INTEGRATION OF AIR QUALITY DATA FOR IMPROVED ESTIMATES OF PM2.5 SOURCE IMPACTS

A Dissertation
Presented to
The Academic Faculty

by

Cesunica E. Ivey

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
August 2016

INTEGRATION OF AIR QUALITY DATA FOR IMPROVED ESTIMATES OF PM2.5 SOURCE IMPACTS

Approved by:

Dr. Armistead G. Russell, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. M. Talat Odman
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. James A. Mulholland
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Rodney Weber
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Yongsheng Chen
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Howard Chang
Department of Biostatistics and Bioinformatics
Emory University

Date Approved: [July 12, 2016]

For Mary, Cecil Jr., and Cecil III

ACKNOWLEDGEMENTS

I sincerely acknowledge the following individuals and groups for their unwavering support during my doctoral studies: Cecil E. Ivey III, my angel on earth; Dr. Tomia B. Hines-Pugh, D.O. and the Hines Family; Odell and Carla Mullen; Dorothy J. Burford; Jennifer D. Burford-Hutchinson and family; Audrey R. Jenkins and family; Prof. Armistead G. Russell and the Russell group students past and present; Dr. Yongtao Hu; Prof. Heather A. Holmes; Ms. Jackie Cox; and Dr. Felicia-Benton Johnson.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Human exposure to air pollution is a known risk factor for the development or worsening of diseases and increased mortality. Traditionally, air quality data available for health analyses are spatially and temporally sparse, which limits the capabilities of the health studies and introduces bias into methods and results. This dissertation is a presentation of novel data assimilation methods using air quality modeling of $PM_{2.5}$ source impacts for use in epidemiologic analyses. The presented methods improve spatially and temporally resolved source impact estimates and address uncertainties in current modeling techniques. The methods developed include a novel hybrid source apportionment method that uses observed and modeled data to generate spatially and temporally resolved source impacts, and impacts of some sources presented here are commonly unresolved by traditional methods (e.g., agricultural/livestock impacts, biogenics, non-road mobile sources, and solvents). The hybrid approach employs a nonlinear optimization method to generate adjustment factors that, when applied to CMAQ-DDM data, revise the original source impacts to better reflect observed concentrations. An initial application and evaluation of this spatial hybrid (SH) approach was conducted for one winter month, January 2004, to estimate daily source impacts over the continental U.S. for 33 source categories. In this application, it was found that biomass burning and dust impacts were subject to the most bias. The SH method was then applied to three additional years (2005-2007). Multi-year trends are discussed, an urban-rural analysis is conducted for 3 U.S. cities, and $PM_{2.5}$ nonattainment was assessed for 3 additional U.S. cities.

Additionally, a novel method for optimizing $PM_{2.5}$ source profiles is presented, which uses observed concentrations to generate a revised profile that reflects local conditions. Revised $PM_{2.5}$ source profiles are generated for monitored locations across the United States, and the spatiotemporal characteristics of the revised profiles are analyzed. Profiles which were dominated by trace metals (K, Fe, Cl, Si) exhibited the greatest variability over space, including coal combustion and metals processing profiles. When using the optimized profiles to calculate concentrations of unspeciated metals, the bias compared to observations decreased compared to the bias in concentrations estimated using the reference profiles. Further, a method is developed to address the modeling uncertainties associated with estimating concentrations and source impacts on secondary $PM_{2.5}$. The method improves the estimates of sulfate, nitrate, ammonium, and secondary organic carbon concentrations, while providing an estimate of numerical bias in source impacts on the species. The sources contributing to the most bias in secondary organic carbon were biogenics, mobile sources, and solvents; agricultural/livestock and mobile sources contributed most to nitrate and ammonium bias; and coal and fuel oil combustion contributed to the sulfate bias.

Overall, the methods and results presented in this dissertation improve the estimates of impacts from the highest contributing $PM_{2.5}$ sources in the U.S., such as vehicle emissions, coal combustion, biomass burning, and dust. Also discussed in this dissertation are the uncertainties in source apportionment and $PM_{2.5}$ modeling techniques, such as limitations in source resolution and spatiotemporal sparseness of results from traditional methods. The methods and results presented generate useful source impact data for scientists interested in policy and health applications, such as establishing

NAAQS attainment for municipalities and further exploring links between human exposure to $PM_{2.5}$ and adverse health effects.

CHAPTER 1

INTRODUCTION

A major anthropogenic pollutant is fine particulate matter (aerosol with an aerodynamic diameter of 2.5 microns or less, $PM_{2.5}$). Ambient $PM_{2.5}$ impacts regional climate by influencing the radiative balance of incoming solar radiation and the outgoing radiation from the earth. $PM_{2.5}$ in high concentrations negatively impacts visibility and high enough concentrations interfere with air travel and grounds flights due to unsafe operating conditions [1, 2]. Most importantly, studies have shown that chronic and acute human exposure to $PM_{2.5}$ has negative impacts on health and quality of life [3-5]. The 2010 Global Burden of Disease Study recently reported that exposure to ambient particulate matter was the ninth leading risk factor of added disability-adjusted life-years (DALYs) globally, and this risk factor was ranked higher than physical inactivity and most dietary deficiencies [6]. One of the first studies to report the association between human exposure to $PM_{2.5}$ and increased mortality was the six-city Harvard study by Dockery et al. (1993), which reported a positive association between death from lung cancer and cardio-pulmonary disease and air pollution [7]. Since then, several air pollution and health analyses were conducted as the seriousness of human exposure to particle pollution became more apparent [8-12].

Recently, the Southeastern Center for Air Pollution and Epidemiology (SCAPE) was formed to further investigate the associations between acute human exposure to air pollution and health outcomes such as cardiovascular- and pulmonary-related emergency department visits, lower birth weights, and childhood asthma [13-16]. This effort was

conducted as a collaboration between the Georgia Institute of Technology and Emory University, where Georgia Tech researchers developed and applied novel approaches for generating high spatial and temporal resolution air quality data and Emory researchers developed and applied novel epidemiologic approaches to explore health correlations. The research presented in this dissertation was motivated by SCAPE objectives and focuses on the improvement of existing and development of new approaches to estimate the impact of pollutant sources on ambient concentrations of $PM_{2.5}$, where the resulting fields are developed for use in health studies (e.g., inputs to epidemiological models) and air quality management (e.g., identification of sources that negatively impact NAAQS attainment designation).

Current source apportionment methods, such as receptor modeling, are limited by spatial and temporal extent, which introduces biases into health analyses that seek to make daily correlations between human exposure and health effects over large study regions (e.g., statewide or regional correlations). Further, daily and spatially complete source impact estimates allow health correlations to be made with specific sources. When important $PM_{2.5}$ sources for health are identified, mitigation and health prevention strategies can be more targeted, since $PM_{2.5}$ concentrations tend to be spatially homogeneous.

The methods and results presented in this dissertation are developed and generated with a data assimilation approach using several sources of data to improve air quality model-derived fields. Other data assimilation approaches for improved fields include a data-fusion method for combining CMAQ concertation fields and observations using a distance-weighting approach [17] and an approach for conducting source

apportionment for organic carbon using a PMF-chemical transport model hybrid approach [18]. In this work, data were integrated to improve source apportionment estimates for gaseous pollutants and $PM_{2.5}$, or the mass partitioning of a pollutant to specific sources of the pollutant (e.g. vehicles, industrial facilities, and residential activities). This dissertation highlights the major sources of $PM_{2.5}$, the spatial and temporal variability of the source impacts, and the uncertainties in modeling source impacts. The studies presented in the chapters that follow are briefly summarized below.

Chapter 2 highlights the development, application, and evaluation of a novel hybrid source apportionment model that provides spatially and temporally resolved $PM_{2.5}$ source impacts for continental U.S. (CONUS). The method further improves upon traditional source apportionment approaches by increasing the number of sources that can be analyzed for one study. In a previous study, it was noted that the choice of the number of sources analyzed may lead to negative contributions, often limiting receptor models [19]. Observations and modeled concentrations of total $PM_{2.5}$ and 40 $PM_{2.5}$ species are assimilated using a nonlinear optimization approach, and the resulting adjustment factors bring modeled air quality fields closer to observations when applied to the original source impact spatial fields. The derivation, application, and evaluation of this hybrid source apportionment at 6 U.S. cities method is detailed in Hu et al. (2014) [20]. The method takes into account the uncertainty in the input data by including numerical uncertainty estimates in the calculations, which is important for health analyses that rely on the generated data. The spatial extension of the hybrid source apportionment method was published in *Geophysical Model Development*, which details the use of geostatistical methods to extend the model over continental U.S. [21]. Optimized source impact spatial

fields are generated for major sources, such as biogenics, mobile sources, coal combustion, biomass burning, sea salt, and agriculture/livestock activities. Optimized source impacts are used to numerically reconstruct concentrations, and new concentrations are evaluated by comparison to observations. Metals concentrations are greatly improved after the optimization step. This application is one of the first instances of the development and evaluation of a method that produces daily, optimized source impacts over CONUS for several unique source categories.

Chapter 3 presents an application of the novel spatiotemporal hybrid method from Chapter 2 for a multi-year analysis. Regional and seasonal trends in $PM_{2.5}$ source impacts are explored for CONUS for nine major $PM_{2.5}$ sources for the years 2005-2007. Results are highlighted for three U.S. regions that were designated as nonattainment in the 2009 $PM_{2.5}$ NAAQS designation for the 2006 standard, and results are compared to previously published attainment studies for the three areas. Seasonal and regional trends for the nine sources of $PM_{2.5}$ are highly variable over the CONUS. This study highlights the usefulness of the novel source apportionment method for regulatory applications, such as establishing $PM_{2.5}$ attainment and identifying the greatest impacting sources for the region with data-assimilated source impacts.

Chapter 4 presents a novel data assimilation method for optimizing $PM_{2.5}$ source profiles, which specify the percent composition of the $PM_{2.5}$ emitted from a source. In this study, profiles for 20 $PM_{2.5}$ source categories and 23 $PM_{2.5}$ species are revised using the novel optimization method, which combines hybrid source impacts and observed concentrations to adjust reference source profiles to better reflect local pollutant conditions. Methods are applied for monitored locations with speciated concentrations

over CONUS for the year 2006, and the resulting optimized source profiles are analyzed for their regional and seasonal characteristics. The performance of the optimized source profiles are evaluated by implementation in a widely-used receptor model. The results from the receptor model application are compared to results from previously-published studies for two U.S. cities. Higher seasonal and spatial variability in revised contributions is seen for metals species that are subject to high bias in modeled concentrations. This source profile optimization method is useful for source apportionment studies in polluted locations where local source profiles are unavailable, and observations can be used to modify reference profiles to better reflect local conditions. The optimization method has been formatted into a user-friendly interface for general use in the air quality community.

Chapter 5 explores the uncertainty of modeling secondary $PM_{2.5}$ by implementing a novel method to improve modeled concentrations and source impacts of sulfate, nitrate, ammonium, and secondary organic carbon (SANOC). In this novel method, the biases in observed and modeled SANOC concentrations are distributed across the source impacts, closing the gap between the model and observations and maintaining the relative contribution of the sources to ambient concentrations of SANOC. Methods are applied over CONUS, and the results are analyzed for their spatial and temporal characteristics. The methods were applied for the year 2006, and this analysis was published in *Frontiers in Environmental Science and Engineering*. The results from this study shed light on the source-specific uncertainties in modeling secondary PM, which can be used for improving future modeling systems. The spatial extent of the method also allows for the

identification of specific U.S. regions where secondary $PM_{2.5}$ modeling performance is subject to the greatest bias.

This dissertation concludes with the major findings from the above studies. Also presented are plans for future research to further improve source apportionment modeling and the implications of the presented research for health and policy studies.

CHAPTER 2

DEVELOPMENT OF $PM_{2.5}$ SOURCE IMPACT SPATIAL FIELDS USING A HYBRID SOURCE APPORTIONMENT AIR QUALITY MODEL

As published in *Geophysical Model Development*

Abstract

An integral part of air quality management is knowledge of the impact of pollutant sources on ambient concentrations of particulate matter (PM). There is also a growing desire to directly use source impact estimates in health studies; however, source impacts cannot be directly measured. Several limitations are inherent in most source apportionment methods motivating the development of a novel hybrid approach that is used to estimate source impacts by combining the capabilities of receptor models (RMs) and chemical transport models (CTMs). The hybrid CTM–RM method calculates adjustment factors to refine the CTM-estimated impact of sources at monitoring sites using pollutant species observations and the results of CTM sensitivity analyses, though it does not directly generate spatial source impact fields. The CTM used here is the Community Multiscale Air Quality (CMAQ) model, and the RM approach is based on the chemical mass balance model (CMB). This work presents a method that utilizes kriging to spatially interpolate source-specific impact adjustment factors to generate revised CTM source impact fields from the CTM–RM method results, and is applied for January 2004 over the continental United States. The kriging step is evaluated using data withholding and by comparing results to data from alternative networks. Data withholding also provides an estimate of method uncertainty. Directly applied (hybrid, HYB) and spatially interpolated (spatial hybrid, SH) hybrid adjustment factors at

withheld observation sites had a correlation coefficient of 0.89, a linear regression slope of 0.83 ± 0.02, and an intercept of 0.14 ± 0.02. Refined source contributions reflect current knowledge of PM emissions (e.g., significant differences in biomass burning impact fields). Concentrations of 19 species and total $PM_{2.5}$ mass were reconstructed for withheld observation sites using HYB and SH adjustment factors. The mean concentrations of total $PM_{2.5}$ at withheld observation sites were 11.7 (± 8.3), 16.3 (± 11), 8.59 (± 4.7), and 9.2 (± 5.7) μg m-3 for the observations, CTM, HYB, and SH predictions, respectively. Correlations improved for concentrations of major ions, including nitrate (CMAQ–DDM (decoupled direct method): 0.404, SH: 0.449), ammonium (CMAQ–DDM: 0.454, SH: 0.492), and sulfate (CMAQ–DDM: 0.706, SH: 0.730). Errors in simulated concentrations of metals were reduced considerably: 295% (CMAQ–DDM) to 139% (SH) for vanadium; and 1340% (CMAQ–DDM) to 326% (SH) for manganese. Errors in simulated concentrations of some metals are expected to remain given the uncertainties in source profiles. Species concentrations were reconstructed using SH results, and the error relative to observed concentrations was greatly reduced as compared to CTM-simulated concentrations. Results demonstrate that the hybrid method along with a spatial extension can be used for large-scale, spatially resolved source apportionment studies where observational data are spatially and temporally limited.

2.1 Introduction

Variations in ambient pollutant species concentrations, including particulate matter (PM) and gases, are correlated with health outcomes–such as lower birth weight [22, 23], higher occurrences of bradycardia and central apnea [24, 25], decreased peak expiratory flows and increased respiratory symptoms in non-smoking asthmatics [26], and all-cause, lung cancer, and cardiopulmonary mortality [27]. Additionally, nanotoxicological studies report that particle uptake by cells and entry into blood and lymph leads to oxidative stress in sensitive areas of the body such as lymph nodes, bone

marrow, and the spleen [28]. Recently, in a study on the global burden of disease, of the 67 risk factors studied, exposure to ambient particulate matter (PM) pollution was the ninth highest risk factor leading to disability-adjusted life years [6]. Many past epidemiologic studies focused on associating PM mass (e.g., $PM2_{.5/10}$: PM with aerodynamic diameters less than 2.5 μm or 10 μm) with the health outcomes, as opposed to individual species or the sources of the PM due to limited data availability or difficulties in quantifying source impacts. Epidemiologic studies are examining the associations between individual species and health outcomes using data from ground observation networks, such as the Chemical Speciation Network (CSN) and the Southeastern Aerosol Research and Characterization Network (SEARCH) [29-32]. It is of further interest to determine the degree to which individual sources are influencing health events and to link human exposure and subsequent adverse impacts to sources and multi-pollutant mixtures [33, 34]. Attributing individual component concentrations and the overall mixture of observed air pollution to specific sources, as well as linking those sources with adverse health impacts is challenging. Typically, receptor models (RM) are used to generate source apportionment (SA) results for epidemiologic studies because longer time series are required (e.g., greater than 2 years) [35].

Several receptor-oriented SA models have been developed to quantify emission source impacts on pollutant concentrations. Each model has its own unique characteristics and associated uncertainties [36, 37]. Schauer and Cass [38] used organic tracers for source apportionment using the chemical mass balance (CMB) method at two urban sites and one background site in central California [39]. Their implementation addressed the improper accounting of volatile organic compounds (VOCs) from motor vehicle exhaust and wood combustion. Watson, Chow, and Fujita [40] reviewed several studies that used CMB for source apportionment, and reported that uncertainties in source contributions of VOCs led to uncertainties in impacts from important sources such as off-road vehicles, solvent use, diesel and gasoline exhaust, meat cooking, and biomass

burning. The authors also describe several limitations of CMB, including reliance on existing observations and overlooking profiles that change between source and receptor due to factors such as dilution, aerosol aging, and deposition. Maykut et al. [41] used positive matrix factorization (PMF) for source apportionment at an urban Seattle, Washington (USA), site with selected trace elements to distinguish combustion sources [42]. Temperature-resolved organic and elemental carbon fractions were also used in Unmix to distinguish diesel and other mobile sources but did not lead to significantly different results [43]. There was also difficulty in distinguishing small sodium-rich industrial sources due to the similarity to the aged marine aerosol source profile.

In an effort to improve the spatial and temporal resolution of SA data and improve source distinction, chemical transport models (CTMs) have been adapted to estimate emission impacts on pollutant concentrations. Marmur et al. conducted a comparison of source-oriented and receptor-oriented modeling results for a winter and summer month in the southeastern U.S. [44]. The brute force method was used in the Community Multiscale Air Quality (CMAQ) model to calculate impacts from mobile sources, biomass burning, coal-fired power plants, and dust. The authors determined that meteorological effects had a strong impact on the temporal variation of CMAQ source impacts, where receptor model results exhibited more day-to-day variability. Koo et al. [45] used the decoupled direct method (DDM) in the comprehensive air quality model with extensions (CAMx) to determine the sensitivity of particle sulfate concentration to changes in emissions of SO2 and NOX from point sources; NOX, VOC, and NH3 from area sources; and all emissions from on-road mobile sources [46-49]. DDM first-order sensitivities underestimated the impacts on sulfate concentration when all emissions are removed due to nonlinearities, as compared to brute force method results. Zhang et al. [50] addressed this issue by calculating second-order sensitivities of inorganic aerosols using DDM, which better captured nonlinear responses to changes in emissions up to 50%.

This work utilizes a hybrid CTM–RM method to provide spatial fields of source impacts for use in detailed health-related, spatiotemporal analyses (e.g., Sarnat et al. 2008) [35]. Spatially resolved source impacts and concentrations are key inputs for residential or county level exposure studies that investigate the impact of air pollution on regional health outcomes [51]. The CTM–RM method combines the strengths of both source apportionment techniques in an effort to reduce uncertainty in source impact estimates. The goal of this study is to create spatial fields of source impacts by spatially interpolating source impact adjustment factors (ratios, or R's) and then applying those adjustments to CTM source impact fields. R's are generated by a hybrid CTM–RM SA approach that integrates observational data and results from a CTM to calculate an emission-based adjustment of source impacts at receptor locations [52]. Kriging is employed to generate spatial fields of R's for 33 emissions sources. The spatial fields of adjustment factors are applied to original source impact fields to produce hybrid-adjusted source impact and species concentration fields for the continental USA The adjustments can also be interpolated in time to adjust source impact fields on days when speciated observations are not available. The performance of the spatial extension is evaluated by performing data withholding and by comparing results to observations from other monitoring networks. The hybrid CTM–RM method, along with the spatial extension, provides air quality data fields for health studies that require spatially resolved exposure metrics. This approach can also be used to assist air quality planners in developing state implementation plans (SIPs) and assessing exceptional events, such as wildland fires.

2.2 Data and Methods

2. 2.1 Data

Observational data from 189 CSN monitors were used for model development and evaluation (Fig. 2.1). Data were obtained on 1 in every 3 or 6 days in January 2004

for a total of 9 days (e.g., Jan. 4th, 7th, 10th…. 28th), which led to varying sample sizes for each observation day. The number of available monitors with speciated $PM_{2.5}$ data on observation days ranged from approximately 40 to 150 and each site had 5 to 9 observations over the period examined. CSN monitor measurements include total $PM_{2.5}$, organic and elemental carbon, ions, and 35 metals. CSN monitors tend to be located in more densely populated areas such as urban and suburban areas, and data are more associated with high-population emissions sources such as mobile and cooking sources. Speciated $PM_{2.5}$ data are also available from the SEARCH [53, 54] and IMPROVE [55] networks, and those data were used for further model evaluation. The SEARCH network includes eight monitors in the southeastern USA, configured as urban/rural pairs. IMPROVE monitors are mainly located in pristine locations such as national parks and wilderness areas. Thirty-eight IMPROVE monitors in the eastern USA were used for model evaluation. IMPROVE monitors in the eastern USA were used due to their closer proximity with urban monitoring sites (e.g., less than 50 km), as opposed to western IMPROVE sites which are much more spatially sparse. Additionally, modeled processes have higher uncertainty for the western USA due to complex terrain and meteorology, leading to added bias in the observation and model comparison [56].

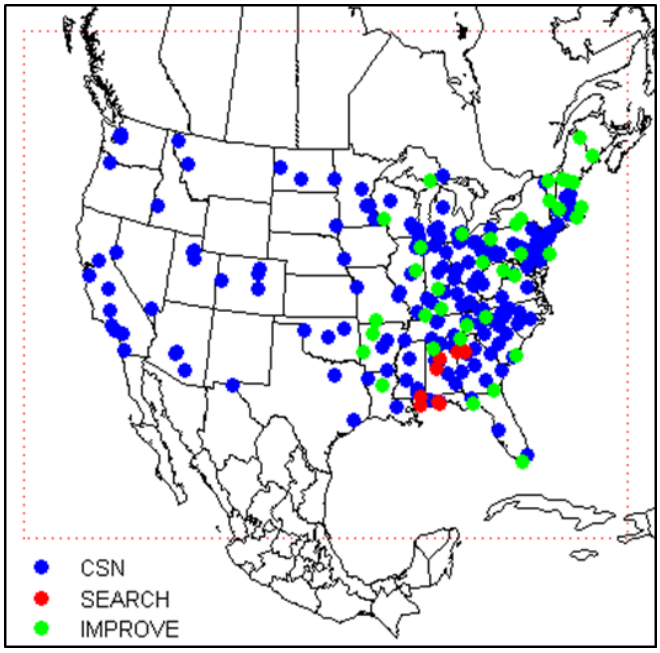

Figure 2.1. Modeling domain (dotted, red line) and CSN, SEARCH, and IMPROVE monitors used for model development, application, and evaluation.

2.2.2 CTM–RM Hybrid Method

This study utilizes a hybrid SA method that combines techniques of both CTMs and RMs to generate adjustment factors (symbolized by R) that improve source impact estimates. Hu et al. [52] described the hybrid approach in detail, but it is briefly summarized here. First, gridded concentrations and emissions sensitivities of $PM_{2.5}$ species are generated using CMAQ–DDM (v. 4.5). CMAQ–DDM model sensitivities to emissions are designated as the original (base case) source impacts ($SA_{i,j}^{base}$) for species *i* and source *j*. CMAQ–DDM was run with strict mass conservation [57], the SAPRC-99 chemical mechanism [58] and the aerosol module described in Binkowski and Roselle [59]. The modeling domain contains the continental USA, southern Canada, and northern Mexico, with 36-km grid resolution, Lambert Conformal Conic geographic projection, and 13 vertical layers of variable thickness extending from the surface to 70 hPa. Meteorological inputs were generated using the fifth-generation PSU/NCAR

(Pennsylvania State University-National Center for Atmospheric Research) mesoscale model (MM5) with 35 vertical layers, implemented with the Pleim–Xiu land surface model [60, 61]. Emissions inputs were processed using the Sparse Matrix Operator Kernel Emissions (SMOKE) model (CEP, 2003). Emissions data originated from a 2004 inventory that was projected from the 2002 National Emissions Inventory (NEI2002). Please refer to the preceding publication by Hu et al (2014) for additional details about the emissions inventory [52].

Next, the original source impacts, receptor observations, and uncertainties are used as inputs to the objective function (Eq. 2.1) of the hybrid SA model:

$$X^2 = \sum_{i=1}^{N}\left[\frac{\left[\left(c_i^{obs} - c_i^{sim} - \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1)\right)\right]^2}{\sigma_{i,obs}{}^2 + \sigma_{i,CTM}{}^2}\right] + \Gamma \sum_{j=1}^{J} \frac{ln(R_j)^2}{\sigma_{\ln(R_j)}{}^2}, \qquad (2.1)$$

where the adjustment factors R_j are optimized by minimizing the objective function, χ^2. The initial R_j values are specific to 1 site and 1 day, as the method is applied at monitors when speciated $PM_{2.5}$ data are available on observation days, and are then kriged and interpolated. The terms c_i^{obs} and c_i^{sim} represent the observed and CMAQ-simulated concentrations, respectively; Γ weights the amount of change in source impact. Uncertainties in observation measurement ($\sigma_{i,obs}$), modeled concentrations ($\sigma_{i,CTM}$), and source strength ($\sigma_{\ln(R_j)}$) are also included in the model. Specifically, $\sigma_{i,obs}$ is reported with measurements for each day from the CSN network; $\sigma_{i,CTM}$ is error in modeled concentrations, which is proportional to observed concentrations and remains constant for all sites and days; and $\sigma_{\ln(R_j)}$ is uncertainty in source contribution expressed as the log of the factor of uncertainty, which also remains constant for each site and day. The uncertainties weight the adjustment of modeled source impacts, in that components with larger uncertainties are weighted less.

The objective function is minimized by using a nonlinear optimization approach known as sequential quadratic programming [62]. The function is modeled using a ridge

regression structure, as demonstrated by the second term, and uses an effective variance approach to balance model outputs. The effective variance approach is also utilized by versions of CMB, and the optimization method used here is, in essence, an extended CMB approach (Watson et al., 1984). Uncertainties in the first term of the objective function serve as effective variances of the numerator and are specified for each species *i*. Finally, R_j are applied to $SA_{i,j}^{base}$ to adjust original source impact estimates (Eq. 2.2) and reconstruct simulated concentrations (c_i^{adj}) at receptors to more closely reflect observations (Eq. 2.3).

$$SA_{i,j}^{adj} = R_j SA_{i,j}^{base} \quad (2.2)$$

$$c_i^{adj} = c_i^{sim} + \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1) \quad (2.3)$$

Given that many of the source impact profiles are similar between categories such that collinearities are present, the variation of the R_j values are constrained to $0.1 \leq R_j \leq 10$.

Source impact profiles are derived from the information provided by Reff et al (2009). In this manuscript, "source impact profiles" are different than "source profiles" in that they describe the source fingerprint at the receptor. In other words, the source profile can be altered, for example by the formation of secondary species. However, for many of the species, there is no secondary formation. It is assumed that within the accumulation mode, which contains most of the fine PM mass in CMAQ, the composition of the primary portion of the $PM_{2.5}$ from any source is the same, but secondary species can be formed, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data are described as follows. Source profiles for 84 source categories were presented in Reff et al.[63], which were aggregated from roughly 300 $PM_{2.5}$ SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 $PM_{2.5}$ profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the contributions in the 33 profiles were used to speciate the "other" (sometimes called unidentified) portion of $PM_{2.5}$ (species name: A25)

as output by CMAQ. The contributions of the 35 trace species were then used to split the "other" $PM_{2.5}$ into individual species, and results for these species, along with the other primary and secondary species are used. At the receptor, both the primary and secondary $PM_{2.5}$ contribution at the receptor is used to determine the new, receptor-oriented, source impact profiles. This same approach was used to generate receptor-oriented profiles in the preceding publication by Hu et al. (2014).

The hybrid method produces results that more closely reflect observations than the original CTM results, which are often biased [20]. It accounts for more known source categories than traditional RM approaches (e.g., 33 versus 6), and it links sources and observations both temporally and spatially. Additionally, the hybrid method generates estimates of the uncertainty in source impact predictions and identifies potential errors in source strength and composition. One limitation of the hybrid method is that results are only available at receptor locations when observations are available, limiting its spatial and temporal scope. In this paper, a spatial hybrid method is presented and evaluated, and it extends the benefits of the hybrid CTM–RM method through spatial interpolation.

2.2.3 Development of spatiotemporal fields

Spatial and temporal source impact fields can be developed by combining the hybrid CTM–RM method and geostatistical techniques. Hybrid-generated Rj values were spatially interpolated for each observation day using kriging to generate spatial fields of source impact adjustment factors. Matlab © (v. 7.14.0.739) was used to perform all geostatistical and optimization calculations. Daily-averaged spatial fields of CMAQ–DDM source impacts are adjusted by grid-by-grid multiplication of the original fields by the corresponding adjustment factor field, resulting in spatial fields of hybrid-adjusted source impacts that are available every third day, as are observations. Source impact fields for intervening periods are developed by interpolation of the Rj spatial fields. Temporally interpolating Rj values and then applying those adjustments to simulated

source impact fields is preferred over simply interpolating the 1-in-3 day hybrid-adjusted source impact fields because temporally interpolating adjusted source impacts would smooth the fields, and the day-specific spatial and temporal variability in the emissions and meteorology captured by the CTM would be lost.

2.2.4 Method evaluation

Performance of the spatial extension was evaluated using a data withholding approach, as well as by comparison with data from the SEARCH and IMPROVE networks. For data withholding, we removed 10% of the available observations (75 sets of observations at the monitors with speciated $PM_{2.5}$ data) and re-ran the spatial hybrid model. This led to a total of 75 observation sets being used in the model evaluation. All references to "withheld CSN data" refer to these 75 sets of withheld data. The remaining 90% of the available observations were used to fit the variogram models, which were used in kriging to produce spatial fields of Rj values. Concentrations are reconstructed using Eq. 2.3 with the spatially interpolated adjustment factors. Additionally, hybrid CTM–RM optimization is directly applied to withheld observation sites to assess the performance of the kriging model. Then the original CMAQ–DDM, directly applied hybrid (CTM–RM), and spatial hybrid (SH) concentrations are compared to measurements at withheld observation locations to evaluate the performance of each method in simulating concentrations. Linear regression was used to assess correlations between observations and modeled concentrations for each method.

In order to evaluate prediction performance in remote locations and in locations independent of CSN, CMAQ–DDM and SH concentrations were compared to observations at SEARCH and IMPROVE locations. Note that the application of the CTM–RM hybrid method, as conducted here, did not include SEARCH and IMPROVE data, and CTM–RM/SH results are independent of those observation data. The SEARCH and IMPROVE comparisons also address the issue of spatial representativeness of using

only CSN data to produce spatial fields. This study uses available speciated CSN data over the entire USA, thereby providing a very spatially heterogeneous data set that is representative of key emissions and meteorology in each U.S. region. The lack of rural data available may present uncertainties in the spatial representativeness of Rj values outside of urban regions.

Also note that 41 species, including total PM, were used for spatial field construction, but only results for 20 species are presented for comparison of CSN results and 15 species for SEARCH and IMPROVE results, as measurements for some trace metals are seldom above measurement detection limit. The possibility of added uncertainty in the optimization step due to detection limit issues was considered. Optimization was tested with the absence of species with limited availability, and no significant differences in model performance were found. The use of the measurement uncertainty in the objective function minimizes the role of those measurements on days when they are below the detection limit, but still accounts for the concentration levels being low. Using all available measurements in the optimization model is the preferred approach.

2.3 Results

2.3.1 Spatial extension evaluation

CTM–RM and SH adjustment factors at withheld observation locations were compared using regression to evaluate the spatial interpolation that was performed using kriging. For each observation day (9 days), 10% of available observations were randomly withheld, resulting in a total of 2,475 R_j data points (75 observations locations x 33 source categories). Five outlying data pairs (< 0.5%) were removed from this regression. Outlying data pairs are determined by examining the distribution of the directly calculated R_j values (mean = 0.84, SD = 0.48) and the kriged R_j values (mean = 0.83, SD

= 0.30) at the withheld observation locations. Data pairs were removed if either value was more than 6 standard deviations from the mean R_j value. The removed data points (5 points out of 2475) were well outside of this range. The remaining CTM–RM and SH factors had a Pearson correlation coefficient of 0.89, a linear regression slope of 0.83 ± 0.02, and an intercept of 0.14 ± 0.02 (Fig. 2.2).

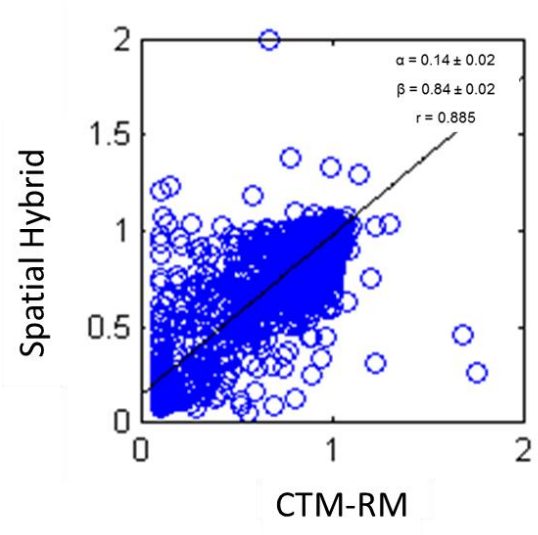

Figure 2.2. CTM–RM vs. spatial hybrid adjustment factors for withheld CSN observations. Regression statistics: intercept, α = 0.14 ± 0.02; slope, β = 0.84 ± 0.02; and correlation coefficient, r = 0.89.

Root mean square errors (RMSEs) were calculated for the adjustment factors by source (Eq. 2.4):

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{N}\left(R_j^{CTM-RM} - R_j^{SH}\right)^2}{N}}, \quad j = 1..J \text{ sources}, \; N = 75 \; sites \qquad (2.4)$$

RMSEs for all sources were less than 0.4, with the exception of RMSEs for lawn waste burning, prescribed burning, and wood stoves (Table A.1). This is expected given the uncertainty in the burn emissions (Table A.2). Sources such as diesel, liquid petroleum gas, non-road natural gas, and Mexican combustion all had very low RMSEs, mean R_j values near 1, and median R_j values near 1. This indicates that there is little to no

adjustment to these source impacts and that kriging captures the R_j values calculated by the CTM–RM application. Mean and median R_j values are within 20% for most sources (Table A.1). The overall mean R_j value at withheld observation locations for all sources for CTM–RM and SH adjustment factors was 0.84 and 0.83, respectively, indicating a high bias in CMAQ–DDM overall, as expected from the base model performance evaluation ($PM_{2.5}$ was biased approximately 40% high).

Cumulative distributions were examined for CTM–RM and SH adjustment factors for each source, and adjustment factors were highly correlated for each source (Fig. A.1). Spatial interpolation captured CTM–RM trends for sources dominated by adjustment factors near 0.1, such as dust, lawn waste burning, prescribed burning, and wood stoves, though did not capture all of the extremely low adjustments (e.g., meat cooking in some locations). Sources that found little adjustment ($R_j = 1$) include aircraft, diesel combustion (stationary sources), fuel oil burning, Mexican combustion, non-road liquid petroleum gasoline combustion, and sea salt, and were well captured by the spatial extension, as demonstrated by nearly identical cumulative distributions. The cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and wood stoves. These sources are highly variable day-to-day, and CMAQ–DDM underestimations are possible in cases where the original emissions missed an actual burn or dust event.

Spatial fields of hybrid adjustment factors are presented for dust, on-road diesel and gasoline combustion, and wood stove sources (Fig. 2.3). Average R_j values over all observation days are also presented for reference (Figure A.2). Typically, R_j values were less than 1 for dust and wood stove impacts, indicating a high bias in those source impacts in the base CMAQ–DDM simulations. Spatial field values for on-road diesel and gasoline combustion R_j are generally near one over most of the USA; however, R_j values for those sources tend be below one in the southeastern region of the USA.

In general, for an R_j value less than 1, the initial CMAQ–DDM estimate is reduced to be more consistent with observations. In turn, for an R_j value greater than 1, the initial CMAQ–DDM estimate is increased to be more consistent with observations. An R_j value of one indicates that no adjustment to the CMAQ–DDM is necessary to improve consistency with observations. As such, after application of the SH method, it was found that while many of the source impacts were adjusted relatively little (i.e., $R_j \approx 1.0$), dust-related and biomass burning-related impacts were often biased high in the original CMAQ–DDM simulation and therefore considerably reduced.

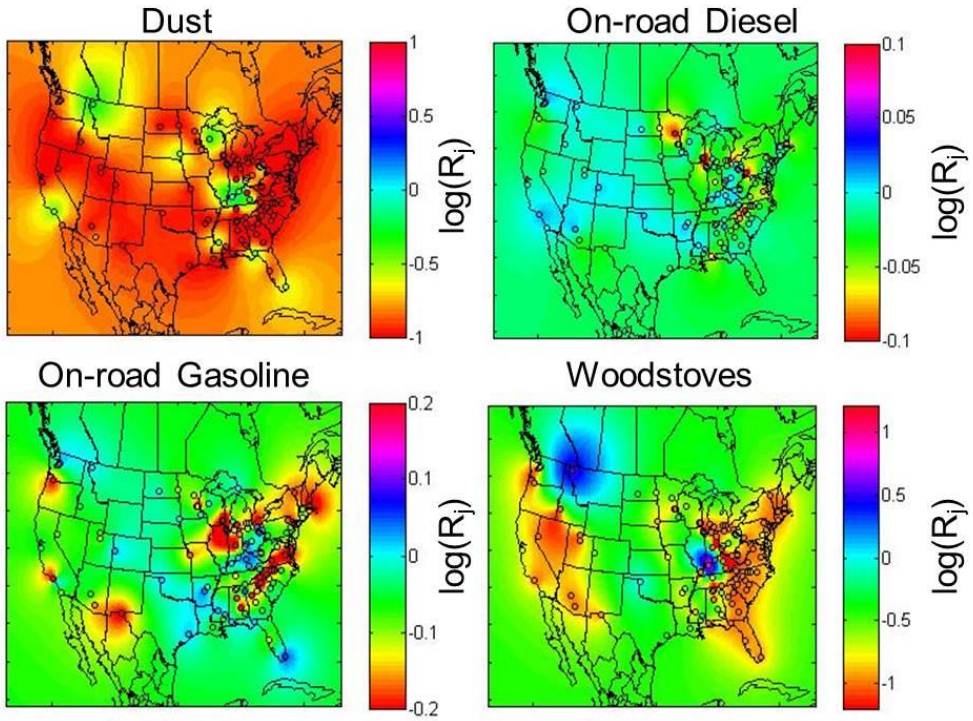

Figure 2.3. Spatial fields of kriged adjustment factors (RjSH) for dust, on-road diesel combustion, on-road gasoline combustion, and wood stove sources for January 4, 2004. Adjustment factors at CSN monitors (denoted by circles) were generated using hybrid (CTM–RM) source apportionment. Note that each panel has a different scale.

The distribution of all R_j values was approximately lognormal, and an analysis was performed to determine whether log-transformation of R_j values prior to the kriging step was necessary to reduce bias in source impact and concentration estimates (Fig. A.3). In one approach, we log-transform the R_j values at the monitors before kriging, and then the kriged values are unlogged before use in reconstruction. In the second approach, we do not log-transform before kriging. From the analysis it was determined that lognormal transformation of R_j values was not necessary, as no significant difference was observed in reconstructed concentrations and source impact fields as a result of the transformation.

Additionally for method evaluation, withheld CSN observations were compared with SH concentrations, which were calculated using kriged R_j values and Eq. 2.3 (Table A.3). The mean concentrations of total $PM_{2.5}$ for withheld observation locations were 11.7 (± 8.3), 16.3 (± 11), 8.59 (± 4.7), and 9.2 (± 5.7) μg m^{-3} for the observations, CMAQ–DDM, CTM–RM, and SH estimations, respectively. Levels of crustal metals (Al, Si, Ca, and Fe), K, and Cl were biased very high in the base CMAQ–DDM simulation, oftentimes an order of magnitude greater than observations. SH concentrations of metals were closer to the CSN observations. Error in simulated (*sim*) concentrations is calculated using Eq. 2.5:

$$Error = \frac{1}{N}\sum_{i=1}^{N}\frac{|obs_i - sim_i|}{obs_i} \quad (2.5)$$

In Eq. 2.5, *i* represents observations and *N* represents the total number of observations withheld for evaluation. The error was 295% and 139% for CMAQ–DDM vs. observations and SH vs. observations, respectively, for vanadium; and 1340% and 326% for CMAQ–DDM vs. observations and SH vs. observations, respectively, for manganese. The large remaining errors stem from the source profiles leading some elements to being biased consistently high and others low. Further work to optimize source profiles can reduce residual errors.

Performance indicators for some species indicate poorer correlation, such as the beta values for calcium for CMAQ–DDM (beta = 1.22) and SH (beta = 0.16) regression comparison (Table A.4). However, all metrics presented must be taken into account and evaluated holistically. The alpha values for calcium indicate an improvement in performance, as the spatial hybrid value (alpha = 0.044) is closer to 0.0 than the CMAQ–DDM value (alpha = 0.13). Further, mean concentrations at withheld observation locations also indicate better performance of the SH model, where mean calcium concentrations were 0.041 (observed), 0.18 (CMAQ–DDM), and 0.050 (SH) (Table A.3). According to the mean concentrations, the SH method performs best. Throughout the analysis, CMAQ–DDM estimates of trace metal concentrations were orders of magnitude too high, while SH results were closer to observations. While some individual metrics indicate better performance of CMAQ–DDM, overall performance of the SH method is most favorable. An important point is that the species where performance is less good are typically those species that have a smaller role in determining source impacts. For example, those species are very trace and/or have high uncertainties in the measurements or source profiles relative to their observed concentrations.

The SH method was further evaluated by comparing simulated concentrations to independent data from the SEARCH and IMPROVE networks (Tables A.5-6). The mean concentrations over observation days were compared, as well as regression statistics for observations versus modeled results. For the SEARCH network (N = 8 monitors), average concentrations of 15 species were compared to observations. Error in mean concentrations for crustal elements was significantly decreased (CMAQ–DDM and SH): Al, 2203 to 540%; Si, 1228 to 271%; K, 365 to 61%; Ca, 402 to 61%; Fe, 260 to 3%; Cu, 231 to 38%; and Se, 63 to 25%. For the IMPROVE network (N = 38 monitors), errors in mean concentrations for crustal elements were also significantly decreased: Al, 704 to 24%; Si, 371 to 24%; K, 599 to 48%; Ca, 361 to 36%; Fe, 334 to 18%; Cu, 186 to 57%; and Se, 22 to 11%. Linear regression metrics are also presented for SEARCH and

IMPROVE monitors (Tables A.7-8). Correlations for all SEARCH and IMPROVE species did not improve; however, estimation performance for most trace metals and ions improved.

2.3.2 Refined source impacts

Refined dust and biomass burning source impacts led to better agreement between simulated and observed concentrations of crustal (Al, Ca, Fe, Si) and biomass burning-derived elements (Cl, K). Original CMAQ–DDM estimates were biased very high for these species compared to observations. This is due to the apparently high bias in source impact profile estimates for biomass burning sources, which do not take into account long-range transport and deposition of biomass burning-related PM. Results suggest that due to atmospheric transformation processes, the source impact profiles are in error for some species, similar to the findings in Balachandran et al. [64]. Observations for some elemental species (Mg, P, V, Se) were highly influenced by measurement limitations (i.e., at or below detection limit) and showed the poorest correlation with modeled concentrations. Additionally, conversion of observed carbon species between analytical methods, from total optical transmittance to total optical reflectance equivalents, introduced potential bias into concentration comparisons. Other studies have shown that conversions may overcorrect observations of carbon species [64].

Average source contributions to $PM_{2.5}$ at withheld CSN observation locations were ranked from largest to smallest for base CMAQ–DDM, CTM–RM, and SH (Table 2.1). The top three sources were wood stoves, dust, and livestock emissions for base CMAQ–DDM simulations, the latter source capturing the influence of ammonia emissions on the formation of nitrate. The livestock category includes impacts from agricultural and farming activities. For CTM–RM and SH results, wood stoves (10^{th} for both) and dust (13^{th} for CTM–RM, 14^{th} for SH) were ranked much lower than for CMAQ–DDM. Livestock emissions were ranked 1^{st} for both the CTM–RM and SH

hybrid applications. Source ranking for open fires was reduced from 10th (CMAQ–DDM) to 20th for both the CTM–RM and SH applications. The fuel oil source impact ranking increased from 12th for the base CMAQ–DDM simulation to 6th and 5th for CTM–RM and SH results, respectively. The order of source contributions at withheld observation locations for the CTM–RM and SH applications compared well, though often differed greatly from the base CMAQ–DDM rankings. The difference in rankings between CTM–RM and SH contributions was, at most, two positions.

The top three sources of primary $PM_{2.5}$ for January 2004, based on source emissions, were dust, wood stoves, and coal combustion, estimated at 1275, 5301, and 3407 metric tons per day, respectively (Table A.2). However, uncertainties associated with dust and wood stove emissions are much higher than most of the other sources, a factor of 10 and 5, respectively [52, 65, 66]. This uncertainty is driven in part by source variability. The large uncertainty and potential bias is reflected in the large shift in rankings for dust and wood stove source contributions to $PM_{2.5}$. Other biomass burning sources such as lawn waste burning and wildfires have similarly large emissions uncertainties and likely large temporal variabilities, and their rankings were also significantly decreased.

Coal combustion includes the secondary formation of sulfate and remains in the top three sources for average SH $PM_{2.5}$ contributions, as its emissions uncertainties are low due to the availability of continuous emission monitoring data. SO_2 emissions are large (Jan. 2004 domain totals: 72924.7 metric tons per day), as are NO_X emissions (74619.7 metric tons per day) (Table A.9). During the study period, coal combustion had the highest contribution to SO_2 emissions (35080.3 metric tons/day) and the second highest contribution to NO_X emissions (14250.1 metric tons per day) behind mobile sources. The source impacts found here account for the transformation of these gaseous emissions from coal combustion.

Secondary formation processes increase the impact of coal combustion, biogenic and livestock emissions relative to their initial primary PM contribution. January 2004 primary PM emissions estimates for biogenic and livestock were ranked 33rd and 31st, respectively. However, CMAQ–DDM, CTM–RM, and SH hybrid contributions ranked both sources significantly higher (biogenic rankings: 14th, 11th, and 9th, respectively; livestock rankings: 3rd, 1st, and 1st, respectively). Although primary $PM_{2.5}$ emissions from these sources are not large, secondary processes and emissions from gaseous precursors led to high source contributions (Table A.9). Biogenic sources emit large quantities of volatile organic compounds which go on to form secondary organic aerosols. Livestock emissions of gaseous ammonia react with sulfate, nitrate, and other acids to form ammonium salts. Therefore, the SH method captures and refines impacts from sources that contribute precursors of $PM_{2.5}$.

Table 2.1. Source category abbreviations with average CMAQ–DDM, CTM–RM, and SH (spatial hybrid) source contributions to $PM_{2.5}$ concentrations for withheld CSN observation locations (N = 75 observations) for January 2004. Note: all averages and standard deviations are expressed in µg m-3. Average total mass of withheld observations, and corresponding CMAQ–DDM, CTM–RM, and SH estimates were 11.7 (± 8.3), 16.3 (± 11), 8.59 ± 4.7, and 9.2 (± 5.7) µg m-3, respectively. NR = Non-road, CM = Combustion.

Source Categories	**Abbreviation**	**CMAQ–DDM**			**CTM–RM**			**SH Hybrid**		
		Avg.	**St. Dev.**	**Rank**	**Avg.**	**St. Dev.**	**Rank**	**Avg.**	**St. Dev.**	**Rank**
Agricultural Burning	AGRIBURN	0.0040	0.003	25	0.0016	0.011	26	0.0012	0.0052	28
Aircraft Emissions	AIRCRAFT	0.0038	0.013	26	0.0037	0.013	25	0.0038	0.013	25
Biogenic Emissions	BIOGENIC	0.074	0.22	14	0.069	0.22	11	0.074	0.22	9
Coal CM	COALCMB	0.16	0.39	9	0.15	0.38	4	0.15	0.38	3
Diesel CM.	DIESELCM	0.00060	0.0017	30	0.0006	0.0017	30	0.0006	0.0017	30
Dust	DUST	0.36	0.095	2	0.061	0.22	13	0.048	0.12	14
Fuel Oil CM	FUELOILC	0.14	0.54	12	0.14	0.62	6	0.14	0.63	5
Livestock Emissions	LIVEST2	0.31	0.89	3	0.31	0.85	1	0.31	0.88	1
Liquid Petroleum Gas CM	LPGCMB	0.0043	0.013	24	0.0043	0.013	24	0.0043	0.013	24
Lawn Waste Burning	LWASTEBU	0.10	0.032	13	0.018	0.067	21	0.010	0.026	22
Metal Processing	MEATALPR	0.18	0.16	7	0.12	0.70	7	0.064	0.22	12
Meat Cooking	MEATCOOK	0.034	0.089	19	0.034	0.10	16	0.032	0.10	17
Mexican CM	MEXCMB_M	0.00070	0.0028	29	0.0007	0.0028	29	0.0007	0.0028	29

Table 2.1 Continued

Source Categories	Abbreviation	CMAQ–DDM			CTM–RM			SH Hybrid		
		Avg.	St. Dev.	Rank	Avg.	St. Dev.	Rank	Avg.	St. Dev.	Rank
Mineral Processing	MINERALP	0.030	0.062	21	0.026	0.075	19	0.024	0.076	19
Natural Gas CM	NAGASCMB	0.17	0.21	8	0.11	0.36	8	0.078	0.20	8
NR Diesel CM	NRDIESEL	0.14	0.48	11	0.14	0.73	5	0.14	0.73	4
NR Fuel Oil CM	NRFUELOI	0.010	0.036	23	0.010	0.041	23	0.010	0.039	23
NR Gasoline CM	NRGASOL	0.063	0.22	16	0.061	0.23	14	0.064	0.23	13
NR Liquid Petroleum Gas CM	NRLPG	0.0014	0.0056	28	0.0014	0.0056	27	0.0014	0.0056	26
NR Natural Gas CM	NRNAGAS	0.0005	0.0014	31	0.0005	0.0014	31	0.0005	0.0014	31
Other NR Sources	NROTHERS	0.0005	0.0012	32	0.0005	0.0012	32	0.0005	0.0012	32
Open Fires	OPENFIRE	0.15	0.099	10	0.021	0.11	20	0.017	0.10	20
On-road Diesel CM	ORDIESEL	0.070	0.17	15	0.066	0.19	12	0.068	0.19	11
On-road Gasoline CM	ORGASOL	0.27	0.60	4	0.20	0.54	2	0.24	0.62	2
Other CM Sources	OTHERCMB	0.040	0.072	18	0.029	0.14	18	0.026	0.11	18
Other PM Sources	OTHERS2	0.18	0.22	6	0.10	0.28	9	0.10	0.28	7

Table 2.1 Continued

Source Categories	Abbreviation	CMAQ–DDM			CTM–RM			SH Hybrid		
		Avg.	St. Dev.	Rank	Avg.	St. Dev.	Rank	Avg.	St. Dev.	Rank
Prescribed Burning	PRESCRBU	0.032	0.054	20	0.031	0.24	17	0.032	0.24	16
Railroad Emissions	RAILROAD	0.013	0.046	22	0.013	0.046	22	0.013	0.045	21
Sea salt	SEASALT	0.0001	0.0005	33	0.0001	0.0005	33	0.00	0.0	33
Solvent Emissions	SOLVENT	0.051	0.094	17	0.044	0.14	15	0.040	0.13	15
Wildfires	WILDFIRE	0.0018	0.0034	27	0.0012	0.0033	28	0.0013	0.00	27
Wood fuel Burning	WOODFUEL	0.22	0.28	5	0.20	1.3	3	0.12	0.90	6
Wood stoves	WOODSTOV	0.62	0.44	1	0.083	0.29	10	0.069	0.28	10

2.3.3 Refined spatial fields

Base CMAQ–DDM spatial fields were refined by applying R_j fields for each source and on each observation day. An example of the adjustment can be found in Figure 2.4, where the CMAQ–DDM spatial field of dust impacts is adjusted on January 4, 2004. Sources with high occurrences (~ >50%) of adjustment factors less than 1 include biomass burning, metals processing, and natural gas combustion, and refined spatial fields for these sources are presented in the supplemental information (SI Figs. A.5-A.7). Biomass burning includes impacts from agricultural burning, lawn waste burning, open fires, prescribed burning, wildfires, wood fuel burning, and wood stoves. The SH method significantly decreases impacts from biomass burning on Jan. 4th and 22nd in the eastern USA and for portions of the west coast (Fig. A.4), largely driven by the observed potassium and organic carbon (OC) levels being lower than simulated levels. On average, CMAQ–DDM simulated levels were a factor of 3.1 (± 1.1) times higher than SH values on Jan. 4th, and a factor of 5.2 (± 1.0) times higher on Jan. 22nd. Metal processing impacts were reduced for areas highly impacted by smelting and metal works industries including the Ohio River valley and mid-Atlantic regions (Fig. A.5). On average, the CMAQ–DDM values were 21 (± 21) % higher than SH values on Jan. 4th, and 25 (± 21) % higher on Jan. 22nd for metal processing impacts. Natural gas combustion impacts (area and point sources only) were reduced for the southeastern USA, the Ohio River valley region, the Gulf of Mexico states, and parts of California and Texas (Fig. A.6). On average, CMAQ–DDM levels were 35 (± 14) % higher than SH values on Jan. 4th, and 72 (± 28) % higher on Jan. 22nd for natural gas combustion impacts.

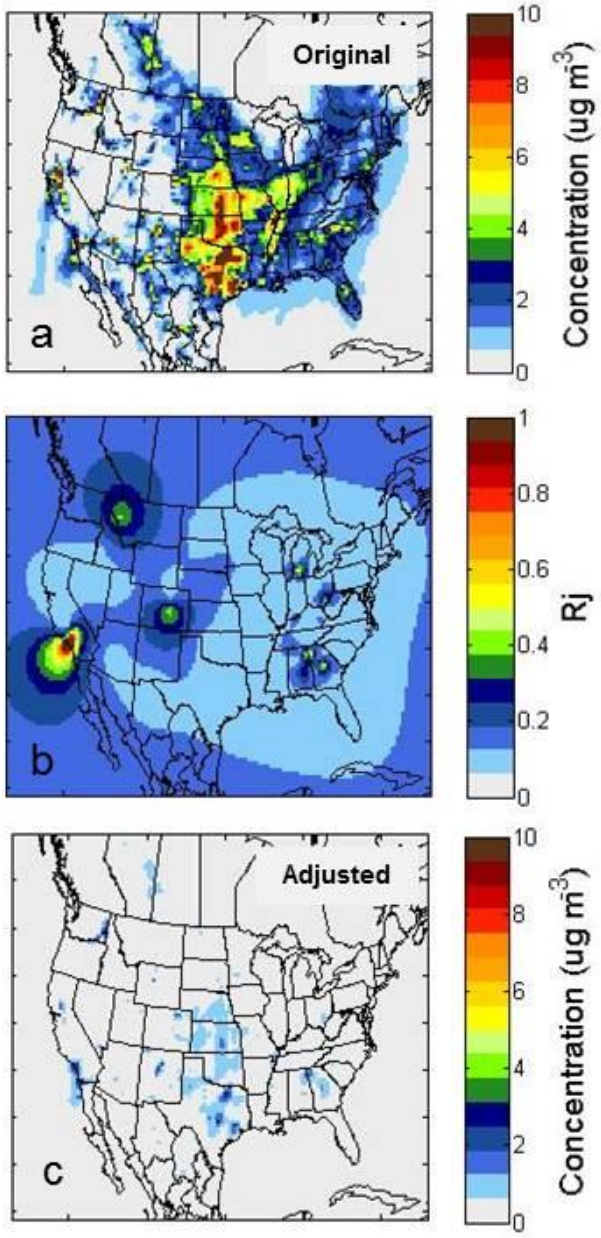

Figure 2.4. Hybrid-kriging adjustment of the dust impacts on $PM_{2.5}$ on January 22, 2004: (a) original CMAQ–DDM simulation of dust source impacts; (b) spatial field of hybrid adjustment factors for dust (RjSH); (c) adjusted spatial field of dust source impacts.

Refined spatial fields of Jan. 2004 averaged source impacts are presented for eight sources: (c,d) dust, (e,f) on-road mobile sources, (g,h) coal combustion, (i,j) sea salt, (k,l) metal-related sources, (m,n) fuel oil combustion, (o,p) biomass burning, and (q,r) agricultural activities (Fig. 2.5). Total $PM_{2.5}$ concentration fields are also included with overlapped observed concentrations from January 28th (a,b). The CMAQ–DDM spatial field overestimates concentrations in the Eastern USA, while overlapped concentrations agree more with spatial hybrid results. Modeled concentrations at monitors in mountainous areas, such as Salt Lake City, Utah, are underestimated due to local meteorological conditions [67, 68]. Wintertime temperature inversions, which cause stagnation in air circulation and consequently high air pollution episodes in industrial valleys, are challenging to capture in models.

Improved spatial field correlation is reflected in monthly averaged spatial fields (Fig 2.5). SH dust impacts are greatly reduced domain-wide as compared to CMAQ–DDM. Monthly averaged refinement of biomass burning, where impacts were also greatly reduced, and metal-related source impact fields are consistent with results previously mentioned for Jan. 4th and Jan. 22nd. Sea salt impacts are localized to coastal areas as expected, and agricultural activity most greatly impacts the mid-western U.S., an area dominated by farm lands. Coal and fuel oil combustion impacts are highest in the eastern USA and western Mexico (fuel oil only) and were adjusted very little as compared to the original CMAQ–DDM field.

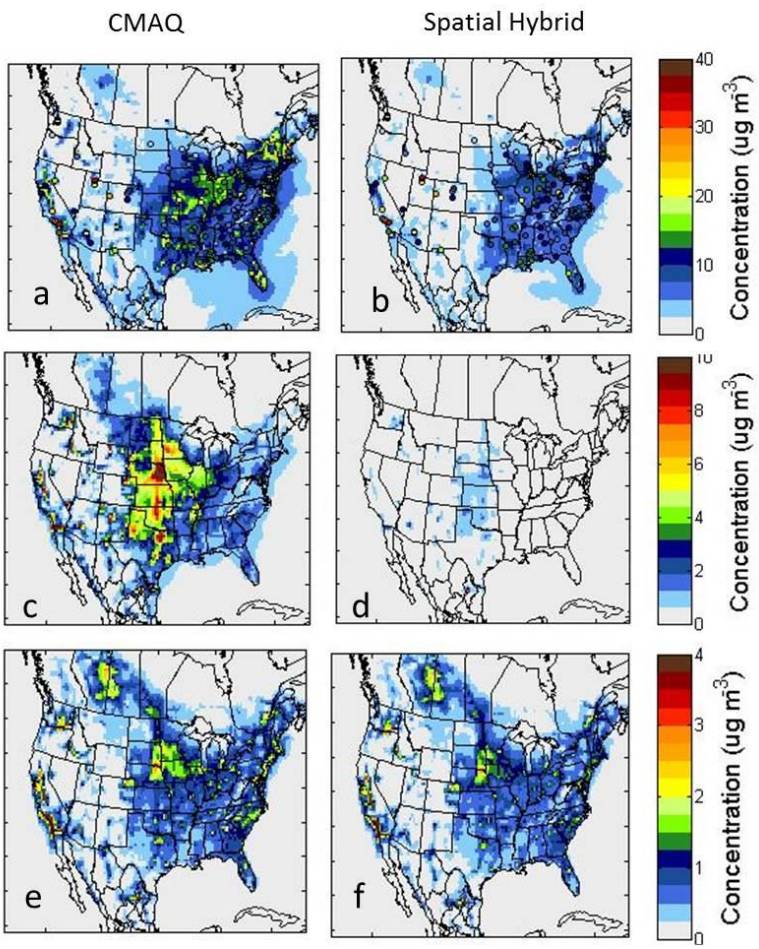

Figure 2.5. Average CMAQ–DDM and spatial hybrid source impacts on $PM_{2.5}$ for observation days in Jan. 2004 for eight source categories. Total $PM_{2.5}$ with overlapped $PM_{2.5}$ observations for Jan. 28th (a,b). Impact of (c,d) soil/crustal material, (e,f) traffic-related sources, (g,h) coal combustion, (i,j) sea salt aerosol, (k,l) metal-related sources, (m,n) fuel oil combustion, (o,p) biomass burning, and agricultural activities (q,r).

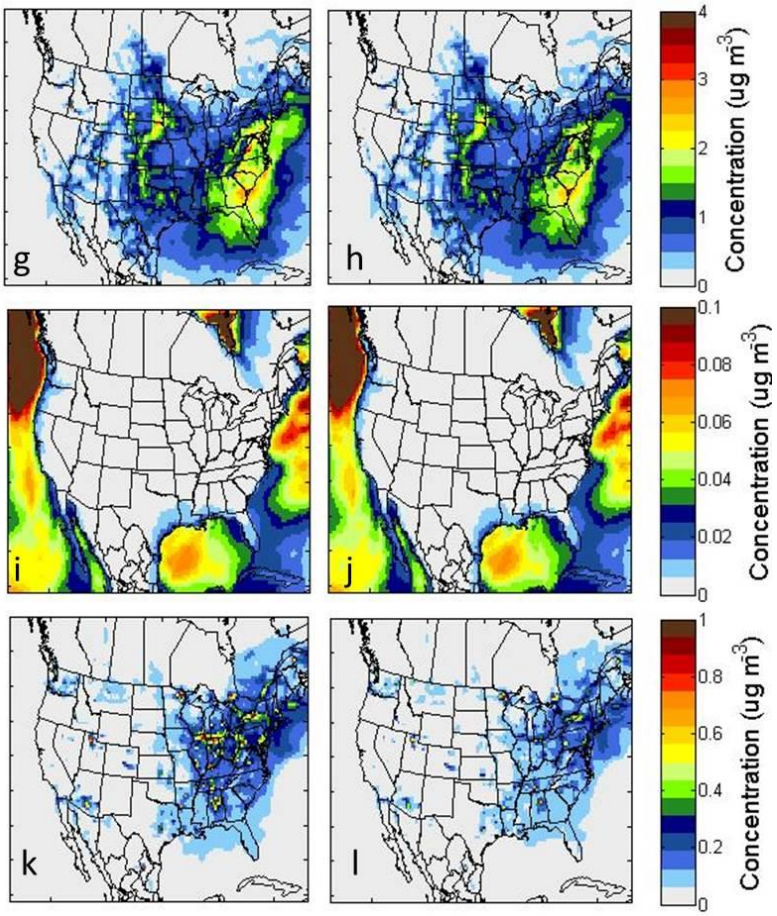

Figure 2.5 Continued.

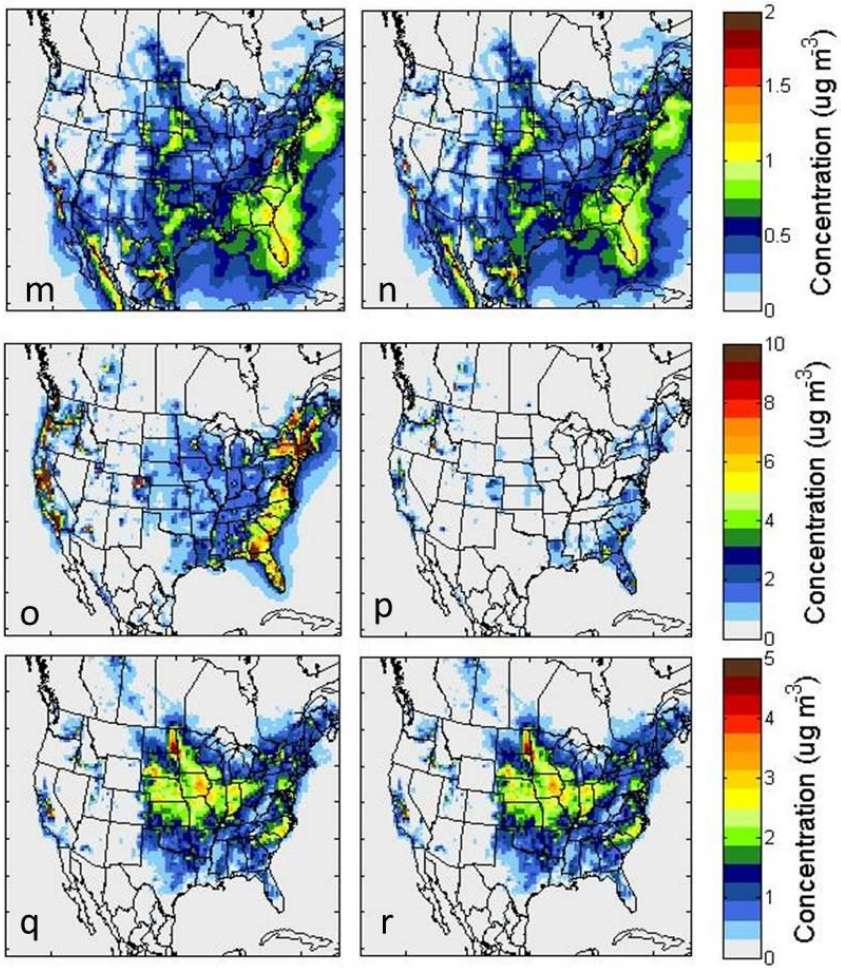

Figure 2.5 Continued.

2.4 Discussion

The SH method uses observations and modeled concentrations of species to adjust impacts on a source-by-source basis to provide spatially and temporally detailed source impact fields. The SH method also captures the impacts of secondary aerosol formation from precursor emission sources. Hybrid adjustment factors can be used to estimate the amount of change in emissions necessary for modeled results to better reflect observations, as emissions are roughly proportional to source impacts for primary sources [52]. Kriging is an effective spatial interpolation method for spatially extending the

CTM–RM model and generating spatial fields of adjustment factors. Kriging does not introduce significant error, as the adjusted fields maintain the spatial and temporal variability of the original fields, and this application led to simulated $PM_{2.5}$ mass concentrations being closer to observations. Adjusted spatial fields of source impacts capture prior knowledge of emissions impacts, meteorology, and chemistry. The SH method also improves simulated estimates of crustal and trace metal concentrations.

The SH method is being developed both to develop spatiotemporally accurate source impact fields that are consistent with observations, and also to provide an approach to increase our understanding of the spatiotemporal characteristics of source impacts in the United States. We find widespread adjustment to biomass burning and dust impacts (Rj less than 1). These source impacts are consistent with observations, emissions estimates, and atmospheric transport and transformation. The SH method is also novel in that, although some sources may not emit a certain pollutant, there still may be some interactions with emissions from other sources leading to those species being part of the source impact. For example, in the case of agricultural fertilizer emissions, although NOx is not directly emitted, the influence on nitrate concentrations is calculated. Although traditionally not quantified in receptor-oriented source apportionment methods, taking into account inter-source interactions is important for determining the primary and secondary impacts of sources on air quality. This hybrid source- and receptor-oriented approach takes this into account and can determine impacts from elusive source interactions. However, this also shows that the formation of secondary species is often dependent upon multiple sources, and the impact of one source is dependent upon other sources, leading to ambiguity in source attribution. The approach here uses the sensitivities at current conditions, though also conducts a mass balance on a species-by-species basis minimizing any overall bias in the source impact attributions.

Spatial hybrid inputs, methods, and results have inherent uncertainties and challenges that are associated with implementation. Input uncertainties include measurement error and challenges are posed with temporal availability and spatial representativeness of concentrations. Emissions inputs for each source are available at different temporal and spatial scales. For instance point source emissions are available at hourly intervals in some cases, while dust emissions are highly variable, both spatially and temporally. Area source emissions are estimated at weekly or monthly intervals and averaged source fingerprints for the primary components of the $PM_{2.5}$ emissions are used, which removes the consideration of locally varying source composition. Physical processes in CMAQ–DDM are uncertain as modeling atmospheric behavior is a complex undertaking. Also, first-order sensitivity approaches may not capture all nonlinearities in source-receptor relationships. SH results are also subject to potential systematic bias from the optimization and kriging steps, though our evaluation suggests those biases are minimal.

2.5 Conclusion

The spatial hybrid model is an effective approach for reducing the error in simulated source impact spatial fields through statistical optimization, instead of re-running CMAQ–DDM which is more computationally expensive. Despite the several points of uncertainty, SH source apportionment can provide daily, spatially complete source impacts across a large domain over a long time period. The SH technique does not necessarily isolate specific atmospheric processes, as it is not a chemistry or physics model. It is a model based on statistics with the assumption that by incorporating observations (truth) and modeled atmospheric processes (prediction), two results can be statistically combined together to yield a better approximation of source impacts. Efforts are continual for reducing uncertainties, increasing the time span of available results, and evaluating estimations with other data sources, such as satellite imagery and independent

field measurements. In future studies, the model will be extended temporally to generate daily, adjusted spatial fields for the continental USA for multiple years and to develop improved source profiles for emissions characterization. Results from SH implementation are beneficial to policy makers, public health analysts, and other air quality scientists that use spatially and temporally complete source impact data in studies where outcomes influence human welfare.

2.6 Acknowledgements

This publication was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901, STAR Fellowship FP-91761401-0**,** and by NASA under grant NNX11AI55G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Alfred P. Sloan Foundation for their support and thank Eric Edgerton of ARA, Inc. for access to the SEARCH data.

CHAPTER 3

REGIONAL AND SEASONAL ANALYSIS FOR A MULTI-YEAR STUDY USING A HYBRID SOURCE APPORTIONMENT MODEL

Abstract

Fine particulate matter is an important criteria pollutant due to the risk of exacerbating existing cardiovascular and respiratory conditions as a result of acute or prolonged exposure. Recent epidemiologic studies have been conducted that seek to determine correlations between $PM_{2.5}$ concentrations and source impacts and the occurrence of adverse health effects. In previous studies, a hybrid source apportionment method was developed that uses chemical transport modeling (CTM) and receptor modeling (RM) frameworks to estimate source impacts. This hybrid CTM-RM method, along with geostatistical methods, is used to generate spatial fields of source impacts that are adjusted to better match observed data. Here, this spatial hybrid (SH) method is implemented for three years 2005-2007 to determine regional and seasonal trends in $PM_{2.5}$ source impacts over the continental U.S. Daily source impacts are generated for 16 unique source categories at 36-km resolution. A comparison is conducted for 3 U.S. metropolitan areas to analyze model representation of impacts on $PM_{2.5}$ in both urban and rural areas. Source impacts are also compared with results from other studies conducted as part of $PM_{2.5}$ NAAQS attainment efforts.

3.1 Introduction

Fine particulate matter pollution, particles with an aerodynamic diameter of 2.5 microns or less, was deemed a criteria pollutant under the Clean Air Act, which was implemented by the U.S Congress in 1970 and amended in 1990. Since the enactment,

the standard for concentrations of primary and secondary $PM_{2.5}$ has changes six times. The initial standard for 24-hour averaged primary $PM_{2.5}$ was 260 μg per m^3 in 1971, then changed to 65 μg per m^3 in 1997, and the current standard set in 2012 is 35 μg per m^3, with other changes in between (U.S. Environmental Protection Agency (EPA)). The tightening of standards over the past four decades is also extended to the annual average standard. The 1971 annual standard was 75 and 60 μg per m^3 for primary and secondary $PM_{2.5}$, respectively. The most recent annual standard, set in 2012, is 12 and 15 μg per m^3 for primary and secondary $PM_{2.5}$, respectively.

The Clean Air Act requires the U.S. EPA to designate areas as either "unclassifiable/attainment" or "nonattainment" after changing or revising the standard (See Supplemental Information for a map of EPA regions). An area is classified as nonattainment if the $PM_{2.5}$ standards have been violated in that area over a three-year period. The EPA issued designations in 2005 for the 1997 standard, in 2009 for the 2006 standard, and designations are currently in progress for the 2012 standard. Many areas moved into attainment after the 2009 designation, even with the tighter standard. Several major metropolitan areas were designated as nonattainment during both periods. In some cases, areas that were deemed in attainment in 2005 were re-classified as nonattainment areas in 2009.

Several source apportionment and $PM_{2.5}$ composition studies were conducted for nonattainment regions in response to the new 2006 standard to determine which sources were affecting the regions and which compounds were dominating the observed $PM_{2.5}$ mass. Pancras et al. conducted a PMF study for measurements collected in July and August 2007 in Dearborn, Michigan, a nonattainment area in the Detroit area [69]. This area in particular has several major industrial facilities such as iron and steel manufacturing, petroleum refining, and lime production. However, it was determined that secondary sulfate and nitrate and mobile sources were the greatest impacting sources,

and not industrial sources. Similarly, Clements et al. conducted a $PM_{2.5}$ composition study (February 2009 to February 2010) for Pinal County, AZ which was designated as a nonattainment area in 2009 for violation of the 2006 standard [70]. The potential sources of $PM_{2.5}$ in this area include railroad and roadway traffic, agricultural and livestock operations, and natural desert emissions. Results from the composition analysis indicated that the dominant sources of $PM_{2.5}$ were crustal materials and organic matter, where seasonality of these species was correlated with tilling and harvesting activities. $PM_{2.5}$ composition studies were conducted for other nonattainment areas such as the San Joaquin Valley, California[71], Fairbanks, Alaska[72], Steubenville, Ohio[73], and Braddock (Pittsburgh), Pennsylvania[74].

The objective of this work is to examine the spatial and temporal trends of $PM_{2.5}$ source impacts over the continental U.S. for the period 2005-2007 with a focus on examining impacts on nonattainment areas and well as identifying the greatest impacting sources and comparing results to previous nonattainment studies. A novel, hybrid source apportionment method with spatial capabilities is employed to generate source impacts over the continental U.S. for 16 distinct source categories, including biomass burning, agricultural activities, livestock operations, and fossil fuel combustion. This study highlights the utility of the spatial hybrid source apportionment methods in evaluating regions for NAAQS attainment.

3.2 Data and Methods

3.2.1 Data

Observations from the Chemical Speciation Network (CSN) are used for model calibration and evaluation (USEPA) (Fig. 3.1). Available from the CSN network is total $PM_{2.5}$ mass, carbon species, major ions, and 35 metal species. $PM_{2.5}$ mass and full speciation measurements were available for 350 of 1,095 (32%) days during the three

year period (2005-2007). Organic and elemental carbon measurements were converted from TOT to TOR equivalents using methods from a previous study [75]. Site-days with any missing species concentrations were excluded from the hybrid optimization.

Figure 3.1: Modeling domain and CSN monitors (blue dots) used in this study.

3.2.2 CMAQ–DDM Modeling

The Community Multi-Scale Air Quality (CMAQ) model is a chemical transport model that is equipped with the ability to estimate model sensitivities to perturbations in inputs or boundary conditions by implementing the decoupled direct method (DDM) for three-dimensional air quality models [46, 49, 76]. CMAQ-DDM modeling was performed for a three-year period from 2005 to 2007, using a 36-km resolution grid applied over the continental U.S. (CONUS), and the modeling domain includes southern Canada and northern Mexico. Meteorological inputs were generated using the Weather Research and Forecasting (WRF) v3.3.1 with the Pleim-Xiu land-surface model, Kain-Fritsch cumulus parameterization, Morrison 2-moment micro physics, RRTM longwave radiation, Dudhia shortwave radiation, and the ACM2 planetary boundary layer scheme [60, 61]. Emissions inputs were generated using the Sparse Matrix Operator Kernel

Emissions (SMOKE) model (v2.6 and v3.1) [77]. The 2005 National Emissions Inventory (NEI) was used for modeling 2005 and 2006 source impacts (SMOKE v2.6). The 2008 NEI was used to produce a 2007-based inventory, and this 2007 inventory was used for modeling 2007 impacts (SMOKE v3.1). Over the course of this study, a new version of CMAQ-DDM was released by the EPA, and two versions were used to model source impacts. The 2006 modeling was performed with CMAQ-DDM v4.7, and the 2005 and 2007 modeling was performed with CMAQ-DDM v5.0.2. The effects of the use of different emissions inventories and modeling versions will be discussed in the results section.

3.2.3 Spatial Hybrid Method

The CMAQ-DDM model was used to generate sensitivities of $PM_{2.5}$ to 16 unique sources (Table 3.1), and these sensitivities are used as the base (or initial) source impact estimates [49]. The spatial hybrid (SH) method for source apportionment combines geostatistical methods with chemical transport model (CTM) and receptor model (RM) capabilities to generate spatial fields of $PM_{2.5}$ source impacts. Details on the SH method can be found in Ivey *et al.* (2015) but methods are briefly summarized here [21]. The hybrid CTM-RM model (Eq. 3.1) is solved using nonlinear optimization, resulting in source impact adjustment factors (R_j) that are used to adjust base case CMAQ-DDM source impacts:

$$X^2 = \sum_{i=1}^{N}\left[\frac{\left[\left(c_i^{obs} - c_i^{sim} - \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1)\right)\right]^2}{\sigma_{i,obs}{}^2 + \sigma_{i,SP}{}^2}\right] + \Gamma \sum_{j=1}^{J} \frac{ln(R_j)^2}{\sigma_{\ln(R_j)}{}^2} \qquad (3.1)$$

In Eq. 3.1, X^2 is the target function to be minimized and is made up of an error term (the first term on the right hand side) and a term that penalizes adjusting source impacts (second term); c_i^{obs} and c_i^{sim} are observed and CMAQ-simulated concentration of species *i*, respectively; $SA_{i,j}^{base}$ is the base-case CMAQ-DDM impact of source *j* on species *i*; $\sigma_{i,obs}$, $\sigma_{i,SP}$, and $\sigma_{\ln(R_j)}$ are uncertainties of the observations, modeled concentrations, and

source impacts, respectively; and Γ is a weighting term to balance the first and second terms and ensure a physically relevant solution. Refined concentrations and source impacts are expressed by Eqs. 3.2 and 3.3:

$$c_i^{refined} = c_i^{sim} + \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1) \quad (3.2)$$

$$SA_{i,j}^{refined} = SA_{i,j}^{base} * R_j \quad (3.3)$$

Kriging is employed to spatially interpolate hybrid adjustment factors (R_j) for each source category at monitoring locations with available speciated $PM_{2.5}$ measurements, where speciation data are typically available every third day. Due to the limited availability of observation data, a temporal interpolation approach is implemented to allow for daily adjustment of source impact spatial fields, instead of adjusting fields only on days with available data. This is achieved by temporally interpolating spatial fields of adjustment factors, where at most two days are interpolated. After grid-by-grid temporal interpolation for each source category, daily spatial fields of adjustment factors are applied to daily base CMAQ-DDM spatial fields. Finally, the SH method yields daily spatial fields of sources impacts that more closely capture observed data.

Table 3.1: Source categories for the multi-year study of source impact spatial fields

Source	**Source**
Agricultural Activities and Livestock Operations	Natural Gas Combustion
Aircraft	Non-road Diesel
Biogenics	Non-road Gasoline
Coal Combustion	On-road Diesel
Dust	On-road Gasoline
Fires	Others
Fuel Oil Combustion	Sea Salt
Metals Processing	Wood Burning

3.3 Results

3.3.1 Multi-year adjustment factors

The spatial hybrid method was applied for years 2005-2007, and distributions of the resulting adjustment factors for 16 source categories are presented in Figure 3.2. The distributions highlight the temporal characteristics of the adjustment factors for each year and each source. The adjustment factors for agriculture/livestock activities, aircraft, coal combustion, fuel oil and natural gas combustion, and non-road diesel and gasoline vehicles are mostly 1, indicating little to no change in the optimized source impact over the three years. For these sources, the 2006 adjustment factors have a slightly larger spread indicating the source impacts in 2006 underwent more change than the 2005 and 2007 impacts. Adjustment factors for biogenic are mostly greater than one for all three years, indicating an overall increase in biogenic impacts after optimization. Adjustment factors for dust, fire, and wood burning have very wide distributions, indicating that the source impact adjustment was variable over all monitors. This is expected due to the highly variable nature of these sources and the high uncertainty in the emissions from these sources. Capturing the spatial and temporal behavior of sporadic wildfires, dust, and residential wood burning is often difficult to capture, and emissions from these sources are typically averaged over a specified time period. For on-road gasoline and diesel sources, distributions are close to one for all three years; however, 2006 had more factors less than one (indicating a reduction in impact) and 2007 had more factors greater than one (indicating an increase in impact). Adjustment factors for metals processing were less than or equal to one for all three years, indicating a reduction in impacts for most locations.

While the general distribution of the adjustment factors is similar over the three year period, there are subtle differences year-to-year that reflect the various biases present in each modeling scenario. The 2006 modeling was conducted using a different version of CMAQ-DDM (v4.7.1) than the 2005 and 2007 modeling (v5.0.2). One major

difference in the modeling versions is that fewer metals were directly output from v4.7.1, which introduces added bias in the simulated metals concentrations, as any metal that is not directly output is derived using source profiles. With inherent uncertainty in the source profiles, the derived metals concentrations are less certain than those directly calculated by CMAQ-DDM. In future work, methods are developed to address the uncertainties in the $PM_{2.5}$ source profiles. (Ch. 4).

3.3.2 Spatial Hybrid $PM_{2.5}$ Concentrations

Modeled concentrations of $PM_{2.5}$ are compared with observations to evaluate the hybrid source apportionment application over three years. Total $PM_{2.5}$ mass at CSN locations was compared to corresponding CMAQ-DDM and SH estimated $PM_{2.5}$. Mean concentrations were as follows: observations (2005: 13.8 ± 8.9 $\mu g/m^3$, 2006: 12.3 ± 7.6 $\mu g/m^3$, 2007: 13.2 ± 8.7 $\mu g/m^3$); CMAQ-DDM (2005: 12.5 ± 8.3 $\mu g/m^3$, 2006: 12.4 ± 8.4 $\mu g/m^3$, 2007: 6.6 ± 5.1 $\mu g/m^3$); and SH (2005: 11.7 ± 7.5 $\mu g/m^3$, 2006: 10.2 ± 6.5 $\mu g/m^3$, 2007: 7.8 ± 8.6 $\mu g/m^3$) (Table B.1). Mean concentrations for 2005 and 2006 CMAQ-DDM were closer to observations than SH, and 2007 SH concentrations were closer to observations. Scatter plots for observations and modeled concentrations of $PM_{2.5}$ indicate little difference in scatter about the 1-to-1 line for all three years (Fig. 3.3). Though mean concentrations from CMAQ-DDM were closer to observations than SH concentrations, correlations between observations and CMAQ-DDM/SH estimates are similar: observations vs. CMAQ-DDM (2005: 0.13, 2006: 0.50, 2007: 0.32); observations vs. SH (2005: 0.13, 2006: 0.55, 2007: 0.36) (Table B.1). Mean bias was higher for SH in 2005 (CMAQ = -1.3, SH = -2.1) and 2006 (CMAQ = 0.0, SH = -2.1) and higher for CMAQ-DDM (CMAQ = -6.7, SH = -5.4) in 2007. These differences may be attributed to the differences in the emissions inventories used for the three modeling scenarios; 2005 and 2006 were modeled using the 2005 emissions platform and 2007 was run using the 2008 emissions platform.

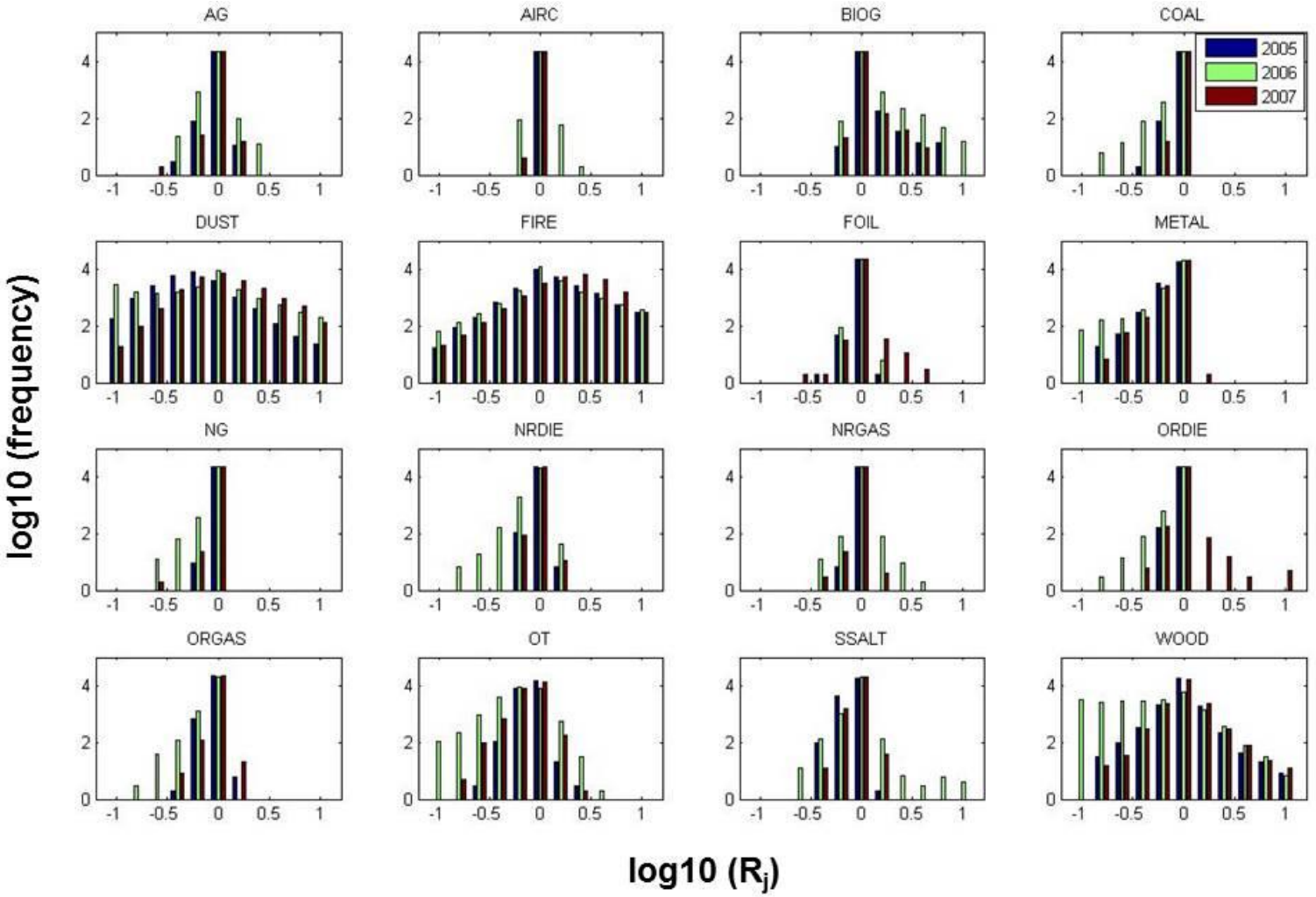

Figure 3.2: Distribution of Rj values for years 2005-2007.

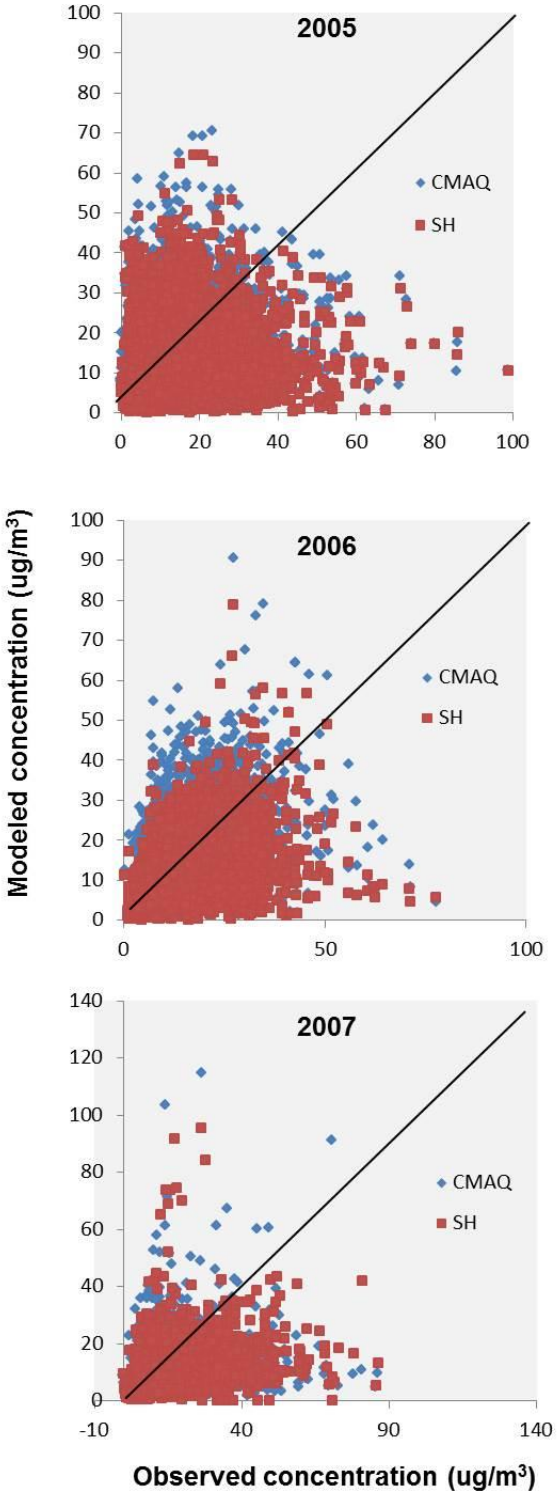

Figure 3.3: Observed and modeled $PM_{2.5}$ scatter plots for 2005-2007. Results are compared for all available CSN monitors.

3.3.3 Multi-year Source Impacts

Hybrid-optimized source impact spatial fields produced for 16 distinct source categories for years 2005-2007 (Table 3.1). Monthly-averaged source impact spatial fields for nine sources are highlighted in this section: agricultural/livestock operations, biogenics, coal combustion, dust, fires, natural gas combustion, on-road diesel vehicles, on-road gasoline vehicles, and wood burning (Figs. B.1-B.9). For agricultural/livestock impacts, peak impacts occur in March and November of 2005-2007 (~5 $\mu g \cdot m^{-3}$), which is consistent with planting and harvest activities (Fig. B.1). The impacts are greatest in the Ohio River Valley region where agricultural activity is highly concentrated. Agricultural/livestock impacts decrease from 2005 to 2007. This occurrence is most likely due to the use of a different algorithm for the temporal allocation of agricultural ammonia emissions in SMOKE v.3.1. The spatial distribution of agricultural impacts is similar over the three years.

For biogenic impacts, peak impacts occur from May to September (~3 $\mu g \cdot m^{-3}$), which is consistent with the active season for most plant life (Fig. B.2). Greatest impacts are seen in the Southeast and the Pacific Northwest. Biogenic impacts are smaller in 2005 and 2007 which is possibly due to the addition of meteorology-based temporal profiles. For coal combustion impacts, the spatial and temporal variability are consistent over the three year period (Fig. B.3). Impacts peak during summer months (~4 $\mu g \cdot m^{-3}$), and display a gradual decrease over the time period of study. The confidence in this decrease as an actual occurrence is high due to the low uncertainty in the emissions from coal combustion sources, as most large sources operate with continuous emissions monitoring (CEM).

For dust, peak impacts are seen in 2007 (~2 $\mu g \cdot m^{-3}$) and are lowest (~0.2 $\mu g \cdot m^{-3}$) in 2006, and impacts vary in magnitude inter-annually (Fig. B.4). Spatially, dust impacts are highest in the central U.S. Dust emissions are processed differently in the two

SMOKE versions where a meteorological adjustment is applied in the 2007 case. For open fire (agricultural, prescribed, and wild fires) impacts, the highest impacts are seen in the southeast in the colder months and western U.S. in the summer months, which is consistent with prescribed burning and wildfire seasons (~5 $\mu g \cdot m^{-3}$) (Fig. B.5). Fire impacts are much higher in 2007 case, likely due to the change in fire emissions from being monthly-averaged in 2005 and 2006 to being provided at daily intervals in 2007. For natural gas impacts, spatial and temporal variability is consistent over the three year period (Fig. B.6). Peak impacts do not follow a seasonal trend (~1 $\mu g \cdot m^{-3}$), and are highest over the eastern U.S. and California.

For on-road diesel vehicles, impacts are highest in 2005 (~2 $\mu g \cdot m^{-3}$) and decrease over the following two years (Fig. B.7). Impacts are highest in the eastern U.S., the Central Valley in California, and in the urban areas of the intermountain west (Salt Lake City, Utah; Boise, Idaho; Santa Fe, New Mexico; Phoenix, Arizona. For on-road gasoline impacts, impacts peak in 2005 and decrease over the next two years (~2 $\mu g \cdot m^{-3}$) (Fig. B.8). There is no apparent seasonal pattern in the on-road gasoline impacts, and impacts are highest in the eastern U.S. and near urban areas. For wood burning impacts (residential and industrial wood burning), highest impacts are seen in the northern half of the U.S. in the colder seasons, which is consistent with wood burning for heating activities (~4 $\mu g \cdot m^{-3}$) (Fig B.9). Higher impacts are seen in 2005 and 2007 because the 2006 optimization reduced wood burning impacts for 2006 more than for other years, possibly due to higher biases in metals concentrations related to wood burning in 2006.

3.3.4 Urban-Rural Analysis

$PM_{2.5}$ concentrations and source impacts are compared for three major U.S. metropolitan areas: Atlanta, GA; Denver, CO; and Los Angeles, CA. Results are presented for the three-year period 2005-2007 at both an urban location and a nearby

rural location to determine the spatial differences in source contributions and model performance compared to observations.

Atlanta, GA

The urban site for Atlanta, GA is located near a major interstate roadway, and the rural site is located in northern GA Mountains. The two sites are located approximately 160 km apart (Fig. 3.4). Overall, for both urban and rural locations, SH concentrations are less than observations where the 3-year averaged $PM_{2.5}$ concentrations were 15.3 (urban obs.), 10.4 (urban SH), 11.4 (rural obs.), and 5.42 (rural SH) $\mu g/m^3$ (Fig. 3.5, top). SH underestimates of $PM_{2.5}$ are more pronounced during winter and fall. The SH model does capture the urban-rural spatial gradient seen in observations where urban concentrations are higher than rural concentrations. The correlation between observations and SH concentrations is better at the urban site (r = 0.47) compared with the correlation at the rural site (r = 0.28). Note that, if either CSN or IMPROVE measurements were missing, data for the day was excluded from the average and correlation calculations.

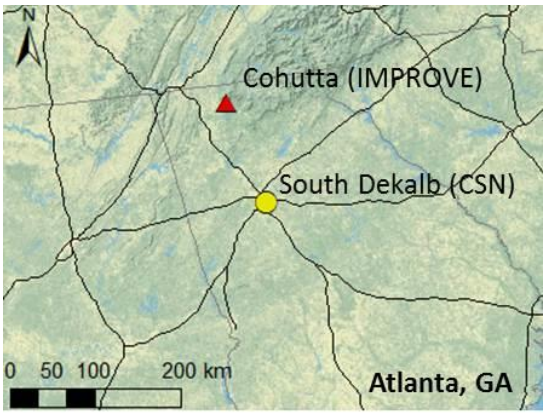

Figure 3.4: An urban and rural monitoring site in the Atlanta, GA area: South DeKalb (CSN) and Cohutta (IMPROVE).

SH source impacts indicate that biomass burning, coal combustion, and mobile sources contribute the largest fractions of $PM_{2.5}$ mass in both the urban and rural locations, with only slight differences between urban (Fig. 3.5, middle) and rural contributions (Fig. 3.5, bottom). For example, mobile sources contribute a higher percentage in the urban location, considering the proximity of Cohutta (rural site) to a major interstate. As stated previously, some year-to-year differences in contributions can be attributed to the use of different emissions inventories for the three-year period, where the 2005 and 2007 results share the common inventory. For example, agricultural impacts contribute a greater percentage of mass in 2006 on average (11%) compared to the other years (2-3%), while biomass burning contributes a greater percentage in years 2005 and 2007 on average (20-40%) compared to 2006 (3%). Coal combustion contribution is slightly lower in 2007 (10% in 2007 and 20% in 2005 and 2006)) and mobile sources contribution is slightly higher in 2006 (30% in 2006 and 20% in 2005 and 2007). Coal and mobile contributions are more stable than biomass burning and agricultural/livestock activities over the three-year period despite the emissions differences.

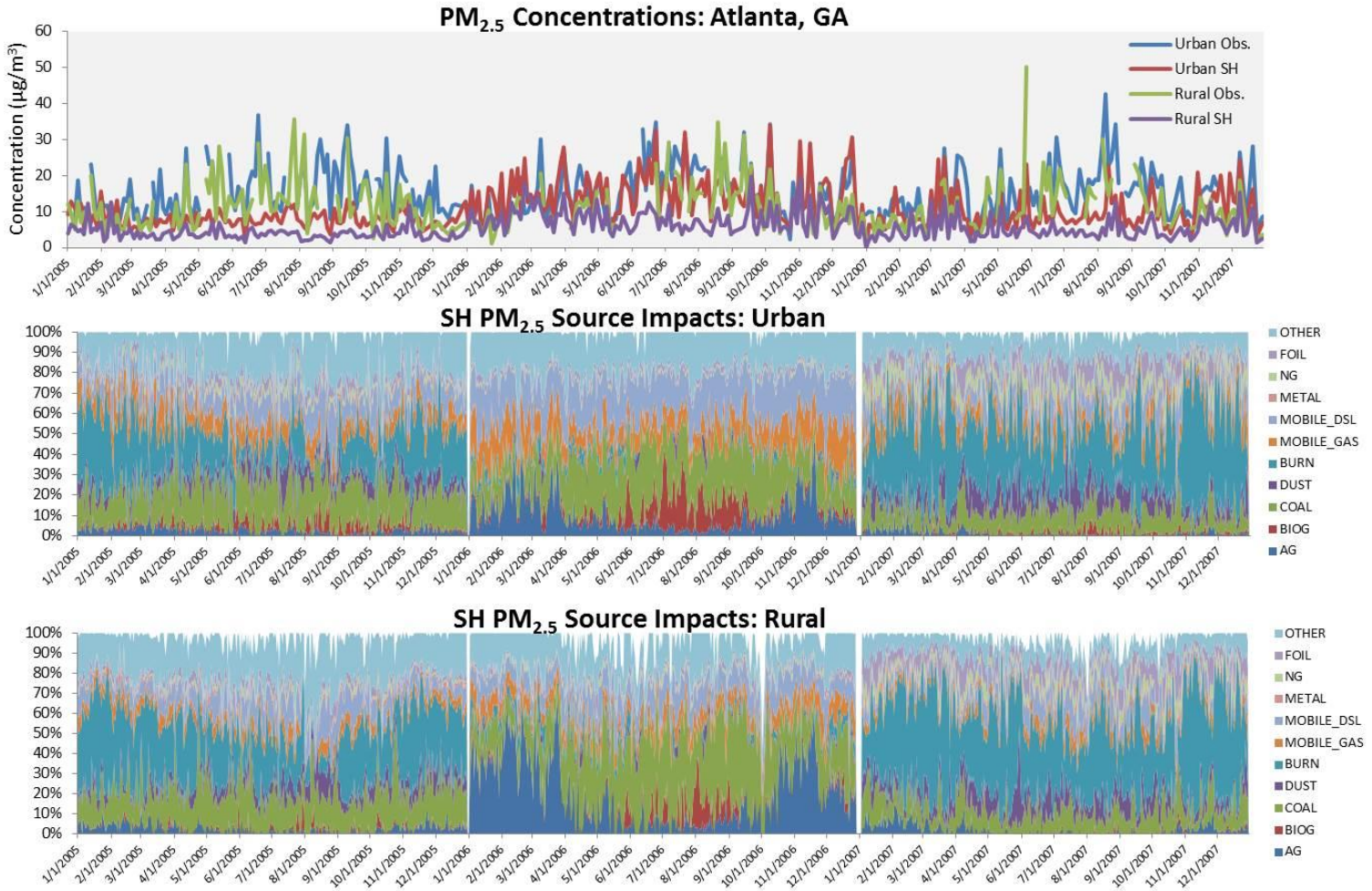

Figure 3.5: $PM_{2.5}$ concentrations ($\mu g/m^3$) and source impacts (%) for the urban and rural monitoring sites in the Atlanta, GA area. Top: Observed and SH modeled $PM_{2.5}$ concentrations for 2005-2007. Middle: SH source impacts for the urban site. Bottom: SH source impacts for the rural site.

Denver, CO

The urban site for Denver, CO is located near a major interstate roadway, and the rural site is located in Rocky Mountain National Park. The two sites are located approximately 100 km apart (Fig. 3.6). Overall, for both urban and rural locations, SH concentrations are slightly less than observations, where the 3-year averaged $PM_{2.5}$ concentrations were 9.55 (urban obs.), 8.26 (urban SH), 3.27 (rural obs.), and 2.99 (rural SH) $\mu g/m^3$ (Fig. 3.7, top). The SH model does capture the urban-rural spatial gradient seen in observations, where urban concentrations are higher than rural concentrations. Correlations between observations and modeled concentrations are not strong (urban: r = 0.15; rural: r = -0.08), indicating that the SH concentrations do not capture daily temporal variations in this location.

Figure 3.6: An urban and rural monitoring site in the Denver, CO area: Birch Street (CSN) and Rocky Mountain National Park (IMPROVE).

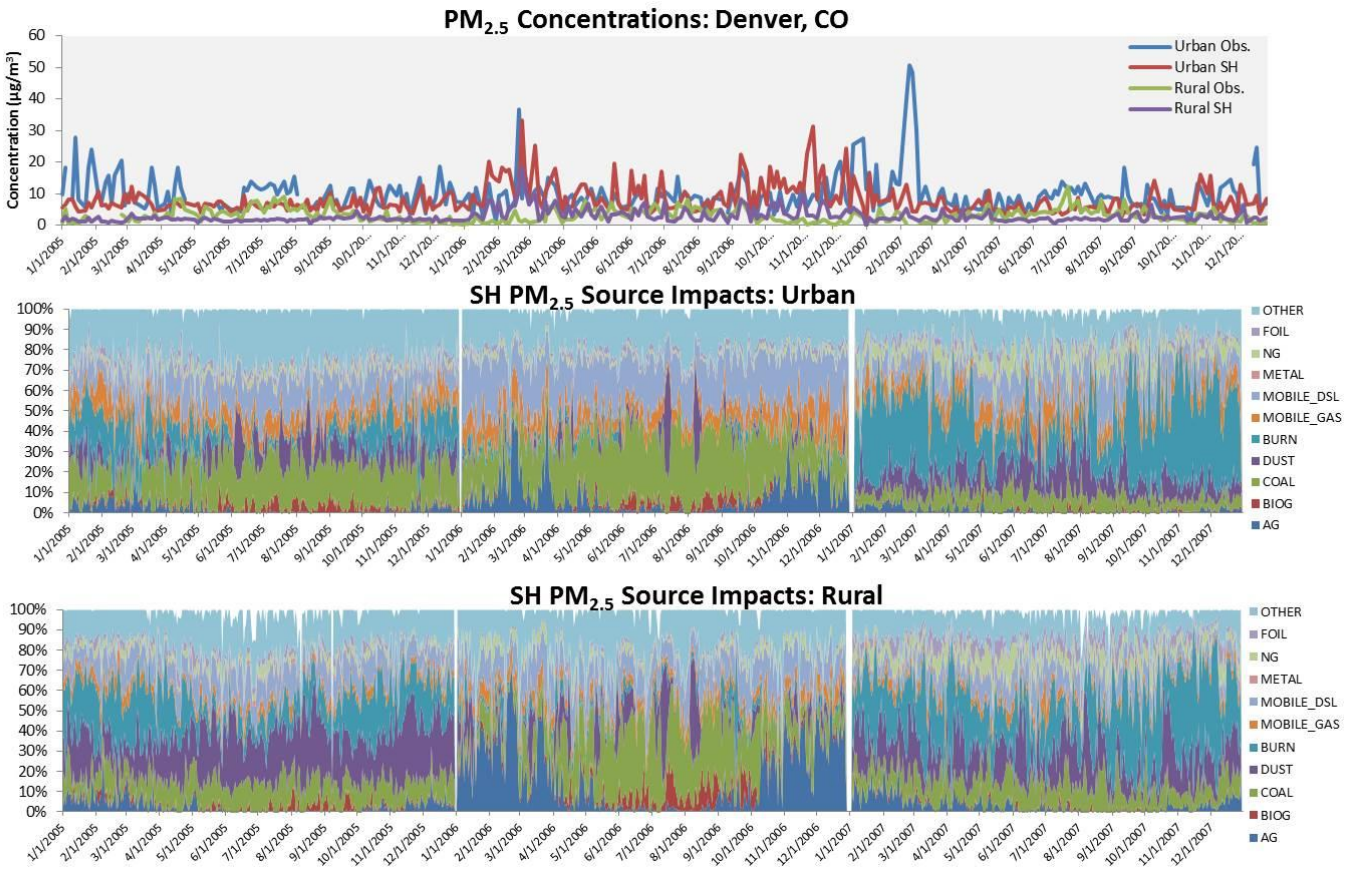

Figure 3.7: $PM_{2.5}$ concentrations (µg/m3) and source impacts (%) for the urban and rural monitoring sites in the Atlanta, GA area. Top: Observed and SH modeled $PM_{2.5}$ concentrations for 2005-2007. Middle: SH source impacts for the urban site. Bottom: SH source impacts for the rural site.

SH source contributions for the Denver, CO area indicate that coal combustion, biomass burning, dust, and mobile sources are the highest contributors in the urban location (Fig. 3.7, middle), while coal combustion, biomass burning, and dust are the highest contributors in the rural location (Fig. 3.7, bottom). The relative contribution of coal combustion is higher at the urban location and is highest in 2006 (urban: 20% (2005), 26% (2006), 8% (2007); rural: 13% (2005), 23% (2006), 12% (2007)). The relative biomass burning contribution is highest in 2007 in both the urban (32%) and rural (29%) locations. The relative dust contribution is higher in the rural location for all three years (urban: 9% (2005), 3% (2006), 12% (2007); rural: 24% (2005), 7% (2006), 17% (2007)). The mobile contribution is less at the rural location, where the average relative gasoline contribution is half of that at the urban location. As seen for the Atlanta area, the highest contribution of agricultural/livestock impacts in the Denver areas are seen in 2006, where the rural contribution is higher than the urban contribution (urban: 8%, rural: 17%).

Los Angeles, CA

The urban site for Los Angeles, CA is located near a major interstate roadway, and the rural site is located in the San Bernardino Mountains. The two sites are located approximately 160 km apart (Fig. 3.8). On average, SH concentrations (9.34 $\mu g/m^3$) were lower than observations (17.2 $\mu g/m^3$) in the urban locations, and SH concentrations (8.44 $\mu g/m^3$) were higher than observations (5.02 $\mu g/m^3$) in the rural location (Fig. 3.9, top). The SH model does not capture the steep urban-rural spatial gradient seen in observations, where urban concentrations are much higher than rural concentrations. Correlations between observations and modeled concentrations are fair and are similar in both locations (urban: $r = 0.33$; rural: $r = 0.32$), indicating that the performance of SH concentrations capturing daily temporal variations is similar in both urban and rural locations. However, the performance in both the urban and rural locations needs improvement.

Figure 3.8: An urban and rural monitoring site in the Los Angeles, CA area: North Main Street (CSN) and San Gorgonio Wilderness (IMPROVE).

The SH source contributions indicate that dominant sources of $PM_{2.5}$ in both urban (Fig. 3.9, middle) and rural (Fig. 3.9, bottom) locations are biomass burning, mobile vehicles, and other (unspecified sources). On average, biomass burning had a higher contribution at the rural site compared to the urban site (urban: 29% (2005), 4% (2006), 30% (2007); rural: 40% (2005), 6% (2006), 33% (2007)). Mobile vehicle contribution was higher at the urban location, and diesel vehicle impacts were higher than gasoline vehicle impacts. Contributions from other (unspecified) sources were similar in the urban and rural locations (urban: 17% (2005), 28% (2006), 12% (2007); rural: 14% (2005), 24% (2006), 15% (2007)). Dust was a more dominant contributor to $PM_{2.5}$ mass in the year 2007.

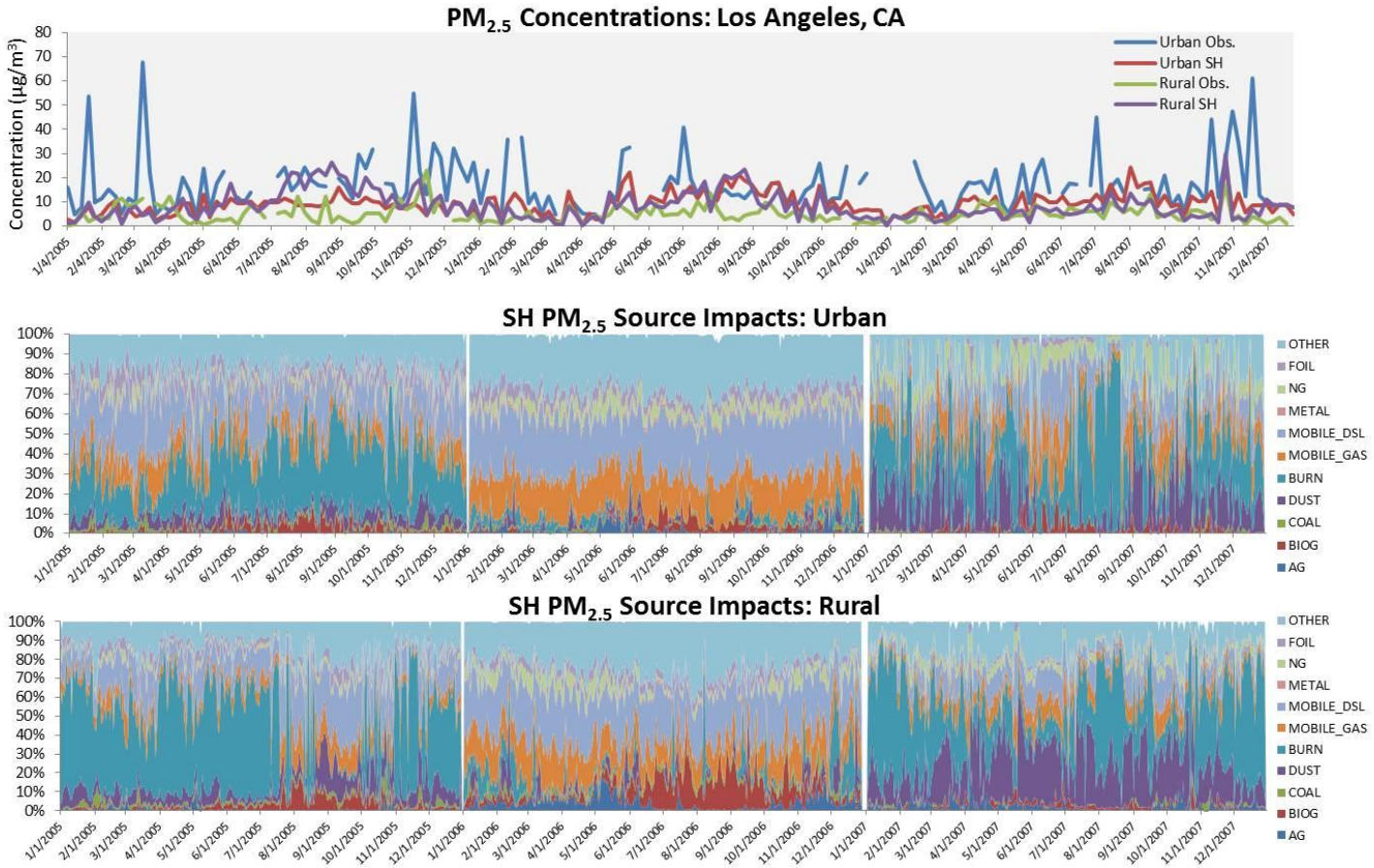

Figure 3.9: $PM_{2.5}$ concentrations (µg/m3) and source impacts (%) for the urban and rural monitoring sites in the Los Angeles, CA area. Top: Observed and SH modeled $PM_{2.5}$ concentrations for 2005-2007. Middle: SH source impacts for the urban site. Bottom: SH source impacts for the rural site.

3.3.5 Receptor Model Comparison

Source impacts are analyzed for three areas designated as nonattainment in 2009 for violation of the 2006 NAAQS for $PM_{2.5.}$ These areas challenged to meet the hourly or annual $PM_{2.5}$ standards due to local emissions and meteorological conditions. Spatial hybrid source impacts are compared to data from receptor modelling studies conducted for these areas during the designation time period. Dearborn, Michigan is a city on the outskirts of the major city of Detroit whose air quality was influenced by several industrial sources (metallurgical coke production, iron and steel production, slag processing, oil refining, electric power generation, automobile manufacturing, metals recycling, incineration, and construction materials facilities) near the time of this current study [69]. Sampling was conducted over 29 days in July and August 2007, where hourly $PM_{2.5}$ mass, gases (NOx, SO2, and CO), elemental carbon, and organic carbon were measured. Source apportionment was conducted using PMF, and results indicated that the highest impacting source was secondary sulfate (6.89 $\mu g \cdot m^{-3}$), followed by road dust/crustal material (1.99 $\mu g \cdot m^{-3}$), secondary nitrate (1.61 $\mu g \cdot m^{-3}$), and mobile sources (1.27 $\mu g \cdot m^{-3}$). Averaged spatial hybrid results for summer 2007 (June, July, August) indicate that the highest impacting sources for Dearborn, MI are fires (2.8 $\mu g \cdot m^{-3}$), other $PM_{2.5}$ sources (1.0 $\mu g \cdot m^{-3}$), coal combustion (0.92 $\mu g \cdot m^{-3}$), fuel oil combustion (0.62 $\mu g \cdot m^{-3}$), and dust (0.53 $\mu g \cdot m^{-3}$) (Fig. 3.10). The spatial hybrid top impacting sources are comparable to the receptor modeling study, as coal combustion emissions are precursors to secondary sulfate formation, and fuel oil combustion is indicative of industrial point sources. The other sources category also includes several industrial point sources. Spatial hybrid dust impacts also appear in the top impacting sources. The state recommendation for nonattainment for Michigan states that the high concentration of industrial sources as the major source of particle pollution [78]. This is consistent with

the SH results, as coal and fuel oil combustion are top sources and are commonly associated with industrial point sources.

Pinal County, Arizona, near the major city Phoenix, was also designated as nonattainment for the 2006 annual $PM_{2.5}$ NAAQS [79]. Sampling was conducted between February 2009 and February 2010 as part of the Desert Southwest Coarse Particulate Matter Study to characterize particle pollution in the area in response to the NAAQS designation [70]. A detailed chemical composition analysis of air and soil samples was conducted to determine possible sources of $PM_{2.5}$. The characterization over three sampling sites found that organic matter is the largest fraction of the fine particle mass (~33%), followed by crustal material (~15%), sulfates (~10%), then nitrates (~7%). Authors suggest that fine PM mass originating from tilling, crop planting, and crop harvesting in spring and fall, which elevates crustal material fractions. Authors also suggest that a nearby cattle feedlot contributes to organic and nitrate fractions. Dust storms are also a contributor of fine particle mass in the area, due to its desert location, as well as pollution from the nearby urban area. Spatial hybrid results indicate that the top impacting sources were diesel sources, dust, and other $PM_{2.5}$ sources in 2005 and 2006, and fires, dust and diesel combustion in 2007 (Fig. 3.11). The SH results are similar to the findings in Clements *et al.* (2014), where dust/livestock is the dominant source of $PM_{2.5}$. The SH results also indicate fires as a dominant source (also a large source of organics) though this source is not identified in the study.

Historically, the San Joaquin Valley (SJV) in California is out of attainment for $PM_{2.5}$ due to the high concentration industrial and agricultural/livestock activity in the region. Also, the mountainous geography of the region impedes dispersion of the pollution, leading to stagnation and temperature inversions that exacerbate air quality issues. Chen et al. (2014) conducted a chemical transport modeling analysis using 2007 and 2019 planning emissions inventories as part of a control strategy to reduce particle

pollution in the region. Measurements taken in the region indicate that ammonium nitrate and carbonaceous compounds dominate $PM_{2.5}$ mass from 2008-2010. Results from the modeling analysis indicate that OC is mainly contributed by diesel equipment, residential wood burning, and meat cooking. Results also indicated that reductions in NOx emissions (on- and off-road mobile sources) would result in the greatest reduction of $PM_{2.5}$, as opposed to reducing ammonia (agricultural/livestock activities) or VOC emissions (residential wood burning). Spatial hybrid results for Fresno County, California (located in the SJV) indicate that agriculture/livestock activities and diesel combustion are the highest impacting sources from 2005-2006, and dust, diesel combustions, biomass burning are the highest impacting sources in 2007 (Fig. 3.12). Results are comparable to those presented in Chen *et al*. (2014), where wood burning, diesel combustion, and agriculture/livestock impacts were most dominant in the SJV region. The SH results do not capture the large agriculture/livestock contribution in 2007, which is likely due to differences in the 2005 and 2008 emissions inventories.

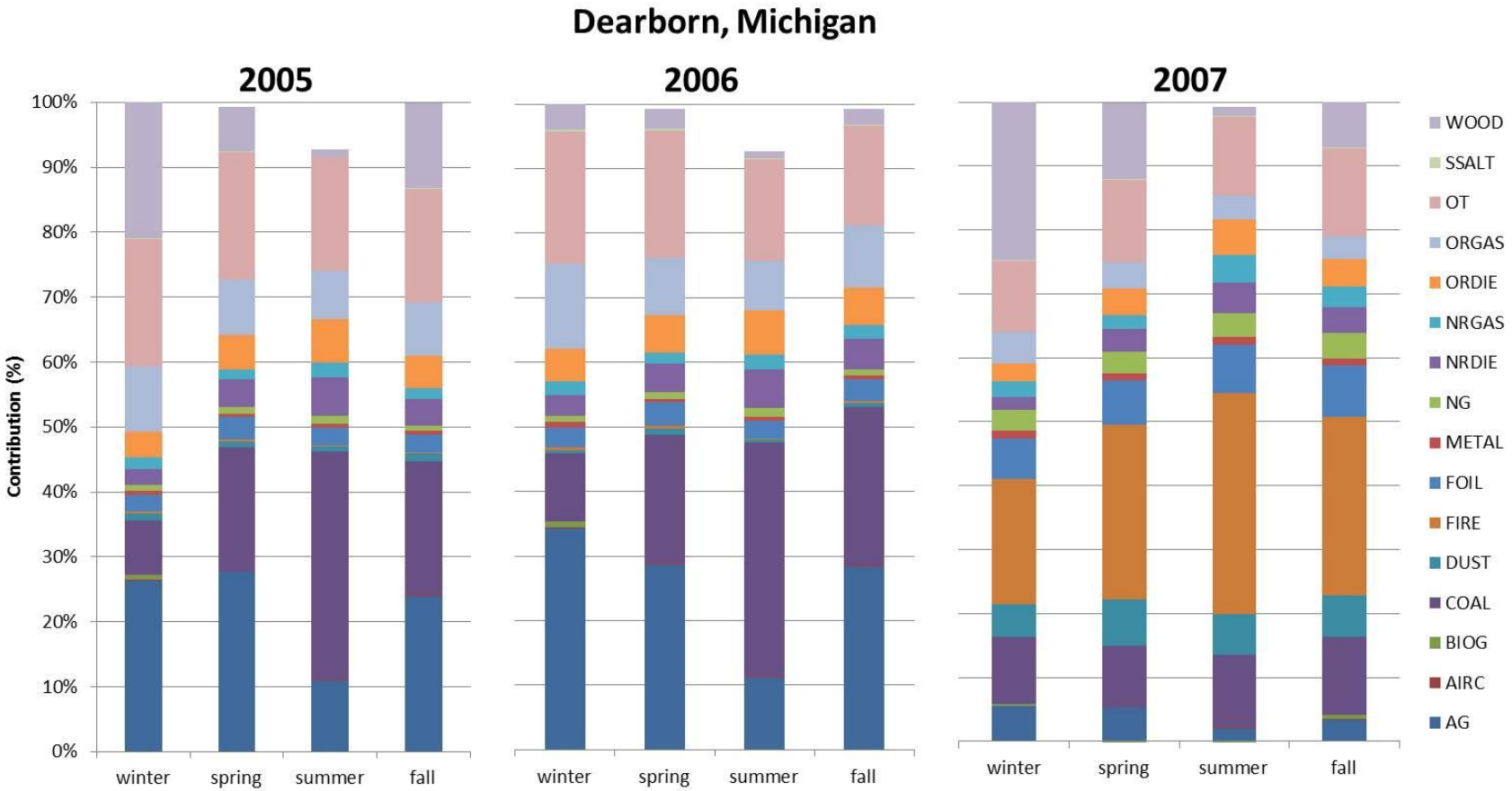

Figure 3.10. Seasonally-averaged spatial hybrid source apportionment results for Dearborn, Michigan for 2005-2007.

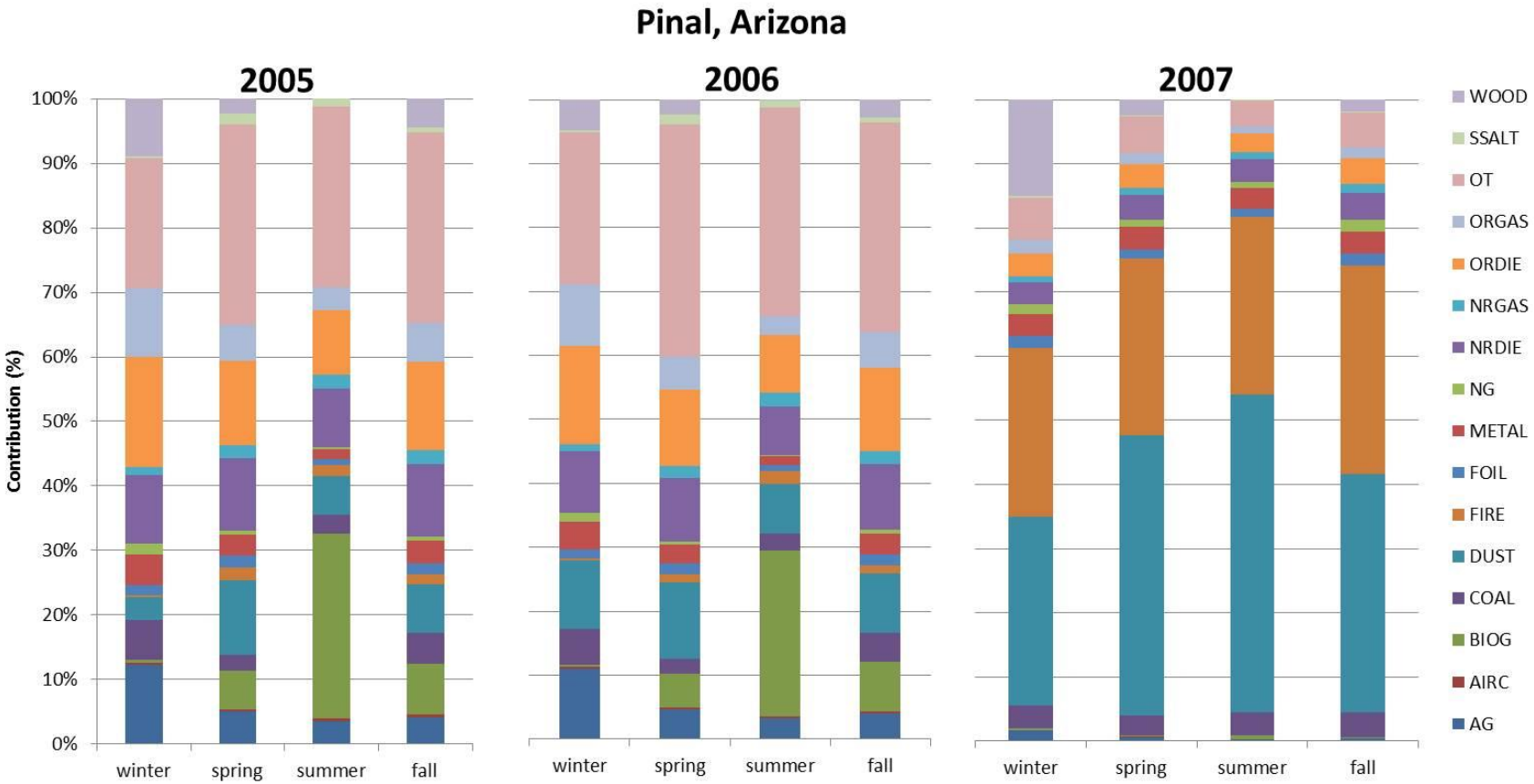

Figure 3.11. Seasonally-averaged spatial hybrid source apportionment results for Pinal, Arizona for 2005-2007.

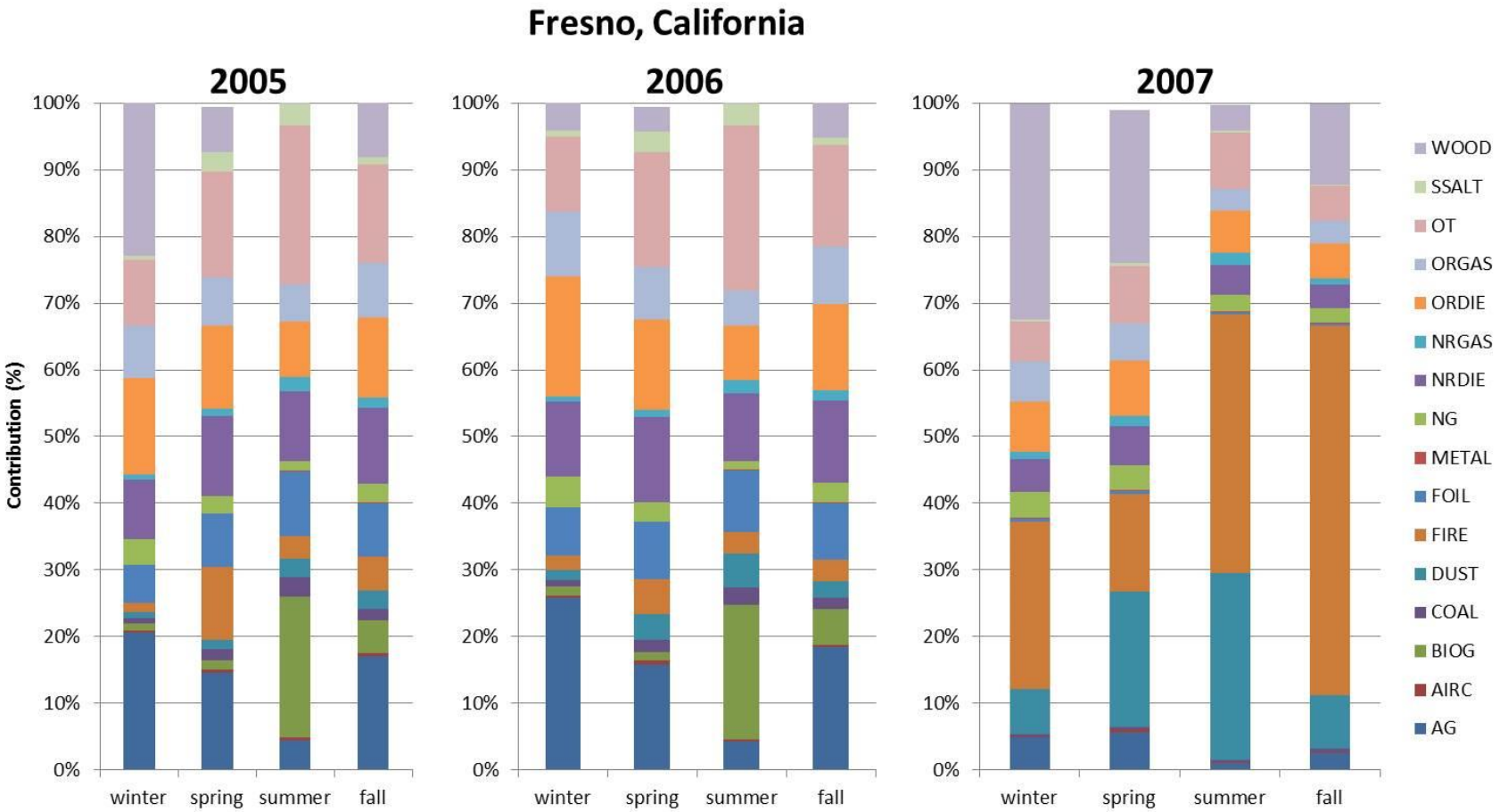

Figure 3.12. Seasonally-averaged spatial hybrid source apportionment results for Fresno, California for 2005-2007.

3.4 Discussion

The source impacts over the U.S. from stationary combustion (coal, fuel oil, and natural gas), mobile sources (on-road and non-road), and aircraft has less variability for the three-year period compared to area sources such as biogenics, dust, and biomass burning (fires and stationary wood burning). Domain-wide surface emissions totals of were calculated for each year to investigate the possible reasons for abrupt changes in area source impacts (Tables 3.2 to 3.4). Some sources in 2005 and 2006 have the same emissions totals because the same inventory was used for both years. The cause of low biogenic impacts in 2007 is not readily apparent when examining emissions, where VOC totals were 263180.0, 266650.0, and 262310.0 metric tons per day for 2005, 2006, and 2007, respectively. The biogenic source impacts for 2005 and 2007 were calculated using the same version of CMAQ, so the cause for discrepancy is not easily identified and requires investigation of day-to-day emissions patterns. Similarly, the 2007 fire source impacts were much higher than in 2005 and 2006. The domain-wide surface VOC emissions from fires were less for 2007 (1139.4 vs 8323.3 (2005 and 2006) metric tons per day). However, the increase in impact for 2007 could be explained by the addition of elevated emissions in the emissions processing, which increases surface concentrations due to vertical mixing. Dust emissions of $PM_{2.5}$ were also similar over the three-year period (19534.0 (2005 and 2006) and 18421.0 (2007) metric tons per day); however the source impacts could be different because of an additional meteorological processing step for the 2007 emissions inventory. Process-level parameterization in CMAQ could be an additional factor in variability of impacts over the period.

Table 3.2: Domain-wide precursor emissions estimates at the surface for 2005 (metric tons/day). Sulfur indicates sulfur trioxide and (SO_3) sulfuric acid (H_2SO_4), POA indicates primary organic aerosol, and PEC indicates primary elemental carbon.

Source Categories	VOCs	$PM_{2.5}$	PM_{10}	SO_2	Sulfur	NO_X	NH_3	CO	POA	PEC	CH_4
Agricultural Activities and Livestock Operations	448.2	1.9	7.4	0.0	0.0	3.6	8276.5	0.0	0.5	0.1	917.9
Aircraft	49.1	12.7	3.2	8.2	0.0	58.0	0.0	591.2	2.7	9.3	3.8
Biogenics	263180.0	0.0	0.0	0.0	0.0	3730.5	0.0	24249.0	0.0	0.0	0.0
Coal Combustion	49.8	297.0	350.6	1431.4	31.2	261.1	6.5	521.4	11.1	8.3	0.0
Dust	185.2	4989.9	19534.0	0.5	0.0	0.9	11.5	3.5	159.9	3.8	10271.0
Fires	8323.3	3230.6	377.8	169.2	0.0	428.6	148.7	26189.0	1726.0	225.8	598.3
Fuel Oil Combustion	40.7	97.5	20.5	1444.0	21.6	457.2	20.8	164.0	18.4	7.7	1.7
Metals Processing	95.4	126.5	29.2	114.1	0.0	24.8	1.9	705.6	6.4	0.9	8.5
Natural Gas Combustion	194.3	35.5	2.7	27.4	0.0	1364.3	28.1	1151.3	10.8	9.8	322.5
Non-road Diesel	1119.8	684.0	25.2	945.3	0.0	7162.8	6.0	3501.0	143.4	502.8	7.1
Non-road Gasoline	11378.0	218.7	15.6	8.1	0.0	427.1	2.3	49768.0	92.7	19.0	903.7
On-road Diesel	357.1	7.1	0.2	3.6	0.0	545.5	0.0	3538.2	3.2	1.7	27.1
On-road Gasoline	1045.2	586.1	56.2	264.2	0.0	7873.6	18.7	3265.7	123.7	396.6	0.0
Others	29810.0	1965.7	331.4	527.9	0.0	3159.8	274.4	10800.4	662.4	100.4	3025.1
Sea Salt	-	-	-	-	-	-	-	-	-	-	-
Wood Burning	1970.1	1417.2	27.5	46.4	0.0	132.4	18.4	7194.0	873.9	54.7	482.8
All Sources	318246.2	13670.3	20781.5	4990.2	52.8	25630.1	8813.8	131642.3	3835.1	1340.7	16569.3

Table 3.3: Domain-wide precursor emissions estimates at the surface for 2006 (metric tons/day). Sulfur indicates sulfur trioxide and (SO_3) sulfuric acid (H_2SO_4), POA indicates primary organic aerosol, and PEC indicates primary elemental carbon.

Source Categories	VOCs	$PM_{2.5}$	PM_{10}	SO_2	Sulfur	NO_X	NH_3	CO	POA	PEC	CH_4
Agricultural Activities and Livestock Operations	448.2	1.9	7.4	0.0	0.0	3.6	8276.5	0.0	0.5	0.1	917.9
Aircraft	49.1	12.7	3.2	8.2	0.0	58.0	0.0	591.2	2.7	9.3	3.8
Biogenics	266650.0	0.0	0.0	0.0	0.0	3822.2	0.0	24439.0	0.0	0.0	0.0
Coal Combustion	49.8	296.8	350.4	1430.9	31.2	261.0	6.5	521.3	11.1	8.3	0.0
Dust	185.2	4989.9	19534.0	0.5	0.0	0.9	11.5	3.5	159.9	3.8	10271.0
Fires	8323.3	3230.6	377.8	169.2	0.0	428.6	148.7	26189.0	1726.0	225.8	598.3
Fuel Oil Combustion	40.7	97.4	20.5	1443.5	21.6	457.1	20.8	164.0	18.4	7.7	1.7
Metals Processing	95.4	126.5	29.2	114.1	0.0	24.8	1.9	705.6	6.4	0.9	8.5
Natural Gas Combustion	194.3	35.5	2.7	27.4	0.0	1364.3	28.1	1151.3	10.8	9.8	322.5
Non-road Diesel	1119.1	683.6	25.2	944.9	0.0	7159.7	6.0	3498.9	143.3	502.5	7.1
Non-road Gasoline	11382.0	218.8	15.6	8.1	0.0	426.8	2.3	49749.0	92.7	19.0	904.2
On-road Diesel	356.6	7.0	0.2	3.6	0.0	544.7	0.0	3533.1	3.2	1.7	27.0
On-road Gasoline	1045.2	586.1	56.2	264.2	0.0	7873.6	18.7	3265.7	123.7	396.6	0.0
Others	29809.5	1965.7	331.4	527.9	0.0	3159.1	274.4	10795.3	662.4	100.4	3025.0
Sea Salt	-	-	-	-	-	-	-	-	-	-	-
Wood Burning	1972.8	1419.0	27.6	46.4	0.0	132.5	18.5	7203.6	875.1	54.8	483.5
All Sources	321721.3	13671.7	20781.4	4988.8	52.8	25716.7	8813.9	131810.4	3836.2	1340.5	16570.4

Table 3.4: Domain-wide precursor emissions estimates at the surface for 2007 (metric tons/day). Sulfur indicates sulfur trioxide and (SO_3) sulfuric acid (H_2SO_4), POA indicates primary organic aerosol, and PEC indicates primary elemental carbon.

Source Categories	VOCs	$PM_{2.5}$	PM_{10}	SO_2	Sulfur	NO_X	NH_3	CO	POA	PEC	CH_4
Agricultural Activities and Livestock Operations	1961.2	810.0	5249.7	41.4	0.0	238.5	10833.0	103.4	38.6	12.8	127.1
Aircraft	119.9	8.7	8.0	20.2	0.0	131.9	0.0	686.0	1.2	3.4	9.0
Biogenics	262310.0	0.0	0.0	0.0	0.0	3645.0	0.0	24079.0	0.0	0.0	0.0
Coal Combustion	14.4	12.4	44.1	311.0	6.2	45.1	0.8	119.7	1.4	0.7	0.3
Dust	248.0	4956.0	18421.0	42.2	0.0	242.1	1335.4	332.1	199.0	19.5	4645.0
Fires	1139.4	999.4	817.3	16.2	0.0	124.4	1.6	4750.4	456.0	71.8	43.1
Fuel Oil Combustion	31.8	53.8	56.5	957.1	14.1	312.9	10.9	101.9	9.3	4.8	11.4
Metals Processing	91.2	142.8	44.4	36.2	0.0	20.2	3.1	277.4	12.0	0.8	3.1
Natural Gas Combustion	205.9	83.0	49.7	79.5	0.0	1242.9	105.3	1058.5	29.2	19.0	218.2
Non-road Diesel	1609.3	578.6	487.5	424.0	0.0	5811.2	5.4	3125.5	121.3	425.2	49.1
Non-road Gasoline	17850.0	219.6	186.2	3.7	0.0	434.7	2.2	45847.0	93.0	19.1	1332.8
On-road Diesel	1010.1	512.9	87.4	24.7	0.0	9342.8	18.2	2910.3	96.6	376.8	3.6
On-road Gasoline	14659.0	286.8	194.5	108.1	0.0	10315.0	395.5	102380.0	146.1	40.6	1027.0
Others	48581.0	1164.0	1256.2	348.5	0.0	2470.9	290.8	5938.8	328.8	35.5	14042.0
Sea Salt	-	-	-	-	-	-	-	-	-	-	-
Wood Burning	3541.8	1545.1	1116.8	27.2	0.0	109.3	44.3	7531.0	960.4	60.0	863.1
All Sources	353373.0	11373.0	28019.3	2440.0	20.2	34486.9	13046.4	199241.1	2492.9	1090.1	22374.9

:

3.5 Summary

The spatial hybrid method was applied for the years 2005 to 2007 to assess changes in seasonal and regional variability in source impacts over continental U.S. Hybrid-adjusted source impacts for 16 source categories were calculated, and spatial fields for nine major sources were highlighted. Coal and natural gas combustion impacts decreased with time while dust and fire impacts increased with time. On-road diesel and gasoline impacts displayed a constant trend over time, and agriculture/livestock, biogenic, and wood burning impacts were highly seasonally influenced. Spatial hybrid source apportionment results were evaluated by citing other studies conducted as part of efforts to improve air quality in nonattainment regions for the $PM_{2.5}$ NAAQS. The other studies identify the top sources that impact the specific region. The top SH sources for the areas analyzed were similar to those presented in the nonattainment studies. This spatial hybrid method is beneficial for attainment applications and identifying the greatest impacting sources on local air quality.

3.6 Acknowledgements

This publication was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901, STAR Fellowship FP-91761401-0**,** and by NASA under grant NNX11AI55G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, the US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Alfred P. Sloan Foundation for their support.

CHAPTER 4

INVERSE MODELING USING A HYBRID CHEMICAL TRANSPORT-RECEPTOR MODELING APPROACH TO DEVELOP REGIONAL $PM_{2.5}$ SOURCE PROFILES

Abstract

Receptor model-based source apportionment studies for fine particulate matter largely rely on source profiles that are estimated from in-situ or laboratory studies that analyze the composition of emissions from the sources in question. In this work, $PM_{2.5}$ source profiles are developed for 20 source categories using data from a novel hybrid source apportionment model. Source profile optimization is conducted for monitored locations for the year 2006, and the resulting profiles are stratified and analyzed by season and EPA region (10 regions in the U.S.). Spatial and seasonal variability is seen for coal combustion, metals processing, and sea salt source profiles over each region when compared to the reference profiles. Optimized source profiles are applied in CMB-iteration to determine the performance of the profiles in traditional receptor models. The results of the CMB-iteration application produced results that were similar to those presented in peer-reviewed literature, supporting the applicability of the hybrid-based profiles. Updated source profiles have the potential to be used in chemical transport models and may reduce the spatial and temporal bias in concentrations of trace $PM_{2.5}$ species that are introduced by source profile errors.

4.1 Introduction

Receptor-oriented modeling is a widely-used approach for estimating the quantitative impacts of particulate matter sources on ambient concentrations. Receptor-oriented techniques largely rely on surface measurements of total $PM_{2.5}$ mass as well as individual $PM_{2.5}$ species. Two of the more popular receptor models are the chemical mass balance (CMB) model and positive matrix factorization (PMF) [39, 80]. The CMB model relies on inputs of $PM_{2.5}$ measurements and estimates of source profiles or "fingerprints"—the mass of individual $PM_{2.5}$ species emitted from a source relative to the total mass emitted [39].

The SPECIATE database is a collection of 5,187 volatile organic gas and particulate matter source profiles developed using data from emission studies and laboratory testing (http://www.epa.gov/ttn/chief/software/speciate/). There are uncertainties associated with profiles in the database, as noted in the development documentation [81]. Profiles for some source classification codes (SCC) are unavailable and are therefore assigned the profile of a comparable source. Profiles developed from laboratory studies with a sample size of N=1 are assigned a lower quality rating in the SPECIATE database, indicating that the profile has higher uncertainty. Additionally, fuel composition may vary by facility (e.g. gasoline refining), hence a composite profile for gasoline vapors is recommended.

Other studies have been conducted to improve $PM_{2.5}$ source profiles. Reff et al. (2009) provide 84 unique source profiles for primary $PM_{2.5}$ which were developed and aggregated using existing profiles from SPECIATE (v4.0) [63]. Authors noted that several profile adjustments were made based on data quality, profile notes, and references associated with the profiles. The 84 profiles include mass fractions of OC, EC, major ions, non-carbon organic matter, metal-bound oxygen, particulate water, 37 metal elements, and other unspeciated $PM_{2.5}$. Sources presented included unpaved road dust,

residential wood combustion, charbroiling, aluminum processing, catalytic cracking, fly ash, phosphate manufacturing, urea fertilizer, potato deep-frying, and steel desulfurization.

In another study, Balachandran et al. (2013) implemented a Bayesian-based ensemble approach for developing season-specific $PM_{2.5}$ source profiles for a winter and summer month for major sources: gasoline vehicles, diesel vehicles, dust, biomass burning, and coal combustion [36]. The ensemble members included PMF, CMB, CMB-LGO (Lipshitz global optimizer), CMB-MM (molecular markers), and CMAQ (Community Multiscale Air Quality model) [46, 82-84]. The source profiles included contributions for 15 $PM_{2.5}$ species, including major ions, carbon species, and trace metals. Certain species in the ensemble-based source profiles for biomass burning and coal combustion showed strong seasonality.

In a study by Lee et al. (2007), the impact of the uncertainties of species measurements and source profile inputs in CMB were quantified using Monte Carlo analysis with Latin hypercube sampling [85]. The percent contributions to uncertainties from individual species in source profiles on CMB results were presented for the following source categories: biomass burning, pulp and paper production, motor vehicles, oil combustion, dust, coal combustion, mineral production, and metal production. Uncertainty contributions of CMB inputs were also presented for ammonium nitrate, ammonium sulfate, and ammonium bisulfate; however, the uncertainties in quantifying impacts from these sources were mainly contributed by the measurement uncertainties for nitrate, ammonium, and sulfate. The findings in the Lee *et al*. study indicate that the uncertainties in the source impacts are highly species dependent.

The objective of this work is to develop new source profiles for 20 sources of fine particulate matter using data assimilation with observations and a hybrid chemical transport-chemical mass balance method. The source profiles are calculated using

nonlinear optimization, and the updated profiles are then used in the CMB-iteration method, which was developed to estimate both primary and secondary organic carbon concentrations, to estimated $PM_{2.5}$ source impacts in two U.S. cities [86]. Optimized source profiles are presented for four seasons and are grouped by their location in the administrative regions of the United States Environment Protection Agency. Regional and seasonal source profile optimization captures spatiotemporal variations in profiles, where static profiles used in traditional modeling may introduce errors by not considering the local variation in emissions.

4.2 Data and Methods

4.2.1 Data

Observations from the Chemical Speciation Network (CSN) are used for method development and evaluation. Available data from the CSN network include total $PM_{2.5}$ mass, organic and elemental carbon, major ions, and 35 metal species. $PM_{2.5}$ mass and full speciation measurements were available for 121 days during the study year (2006), as measurements are available every third or sixth day depending on the local monitoring schedule. The number of monitors available on observation days varied from 40 to 150 due to the previously stated scheduling. Carbon species concentrations were converted from TOT to TOR equivalents using methods from a previous study [75]. In this study, only 23 species are used for source profile development due to the low occurrence of concentrations below detection limit for these species [75].

4.2.2 CMAQ–DDM Modeling

The CMAQ model is a chemical transport model that outputs gridded concentrations of atmospheric pollutants and is equipped with the ability to estimate model sensitivities to perturbations in inputs or boundary conditions by implementing the decoupled direct method (DDM) for three-dimensional domains [49, 76]. CMAQ-DDM

modeling was performed for one year 2006 at 36-km resolution over the continental U.S., and the modeling domain includes southern Canada and northern Mexico. Meteorological inputs were generated using the Weather Research and Forecasting (WRF) model v3.3.1 with the Pleim-Xiu land-surface model, Kain-Fritsch cumulus parameterization, Morrison 2-moment micro physics, RRTM longwave radiation, Dudhia shortwave radiation, and the ACM2 planetary boundary layer scheme [60, 61]. Emissions inputs were generated using the Sparse Matrix Operator Kernel Emissions (SMOKE) model (v2.6) along with the 2005 U.S. EPA National Emissions Inventory [77]. The CMAQ-DDM model was used to generate $PM_{2.5}$ sensitivities to 20 unique sources, and these sensitivities are taken to be the equivalent to base or initial source impact estimates (Table 4.1).

Table 4.1. Source categories and abbreviations used in the source profile study.

source	**abbrev.**	**Source**	**abbrev.**
agricultural activities and livestock operations	ag	non-road diesel	nrd
aircraft	air	non-road gasoline	nrg
biogenics	biog	non-road others	nro
coal combustion	coal	on-road diesel	ord
dust	dust	on-road gasoline	org
fires (wildfires, prescribed burns)	fire	Others	ot
fuel oil combustion	foil	other combustion	otc
meat cooking	meat	Solvents	slv
metals processing	metal	sea salt	ss
natural gas combustion	ng	wood burning	wood

4.2.3 Spatial Hybrid Source Apportionment

A hybrid chemical-transport-receptor (CTM-RM) model was developed that uses base case CMAQ-DDM estimates and observed concentrations to optimize modeled source impact estimates. The approach is detailed in Hu et al. (2014) [20], and briefly

described here. The hybrid CTM-RM source apportionment model is applied at CSN monitors with speciated $PM_{2.5}$ data. The model optimizes an adjustment factor R_j (for source *j*), which is applied to base case CMAQ-DDM source impacts to either increase or decrease the impacts so that modeled estimates better reflect observed concentrations (Eq. 4.1).

$$X^2 = \sum_{i=1}^{N}\left[\frac{\left[\left(c_i^{obs} - c_i^{sim} - \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1)\right)\right]^2}{\sigma_{i,obs}{}^2 + \sigma_{i,SP}{}^2}\right] + \Gamma \sum_{j=1}^{J} \frac{ln(R_j)^2}{\sigma_{\ln(R_j)}{}^2} \quad (4.1)$$

In Eq. 4.1, X^2 is the error to be minimized; c_i^{obs} and c_i^{sim} are observed and CMAQ-simulated concentrations of species *i*, respectively; $SA_{i,j}^{base}$ is the base-case CMAQ-DDM impact of source *j* on species *i*; $\sigma_{i,obs}$, $\sigma_{i,SP}$, and $\sigma_{\ln(R_j)}$ are uncertainties in the observations, modeled concentrations, and source impacts, respectively; and Γ is a weighting term to balance the optimized output and ensure a physically relevant solution. Refined concentrations and source impacts are expressed by Eqs. 4.2 and 4.3:

$$c_i^{refined} = c_i^{sim} + \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1) \quad (4.2)$$

$$SA_{i,j}^{refined} = SA_{i,j}^{base} * R_j \quad (4.3)$$

The hybrid CTM-RM method is extended spatially by kriging adjustment factors R_j and gridded R_j values are applied to the corresponding gridded $SA_{i,j}^{base}$ spatial fields [21]. The gridded R_j values, originally available only on observation days, are temporally interpolated using grid-by-grid linear interpolation to generate daily, 36-km resolution spatial fields of refined source impacts and concentrations across the contiguous U.S.

4.2.4 Source Profile Optimization

Source profiles are developed for the 20 $PM_{2.5}$ sources using a nonlinear optimization approach that uses CSN observations and CMAQ-DDM source impacts to generate receptor-based profiles. The optimization equation is applied at one monitor for

one observation day at a time. We derive the source profile optimization equation (Eq. 4.9) from the CMB method:

$$c_i^{obs} = \sum_{j=1}^{J} f_{i,j} S_j + e_i \tag{4.4}$$

where $f_{i,j}$ is the fraction of species i that is emitted from source j, S_j is the impact of source j on total $PM_{2.5}$, and e_i is the concentration prediction error to be minimized. Similarly, for the spatial hybrid impact of source j on total $PM_{2.5}$, the following equality holds:

$$S_j^{hyb} = R_j \sum_{i=1}^{I} SA_{i,j}^{base} \tag{4.5}$$

under the assumption that all impacting sources on total $PM_{2.5}$ are accounted for by the spatial hybrid approach. To calculate new source profiles, the variable $f_{i,j}$ and is expressed as the product of an adjustment ratio and the original or reference species fraction, $f_{i,j}^0$:

$$f_{i,j} = r_{i,j} f_{i,j}^0 \tag{4.6}$$

Now the variable of interest becomes $r_{i,j}$ (upper bound = $2 * r_{ij}$ and lower bound = 0). The new equality becomes:

$$c_i^{obs} = \sum_{j=1}^{J} r_{i,j} f_{i,j}^0 (R_j \sum_{i=1}^{I} SA_{i,j}^{base}) + e_i \tag{4.7}$$

The source profile ratio is optimized by minimizing the square prediction error, X^2:

$$X^2 = \sum_{i=1}^{I} c_i^{obs} - \left(\sum_{j=1}^{J} r_{i,j} f_{i,j}^0 (R_j \sum_{i=1}^{I} SA_{i,j}^{base}) \right) \tag{4.8}$$

Uncertainties and constraints are added to the objective function to weight and balance the prediction error:

$$X^2 = \sum_{i=1}^{I} \frac{\left[c_i^{obs} - \left(\sum_{j=1}^{J} r_{i,j} f_{i,j}^0 \left(R_j \sum_{i=1}^{I} SA_{i,j}^{base}\right)\right)\right]^2}{\sigma_i^{obs\,2} + \sigma_i^{SA\,2}} + \alpha \Gamma \sum_{i=1}^{I} \left[\sum_{j=1}^{J} \frac{f_{i,j}^0 (r_{i,j} - 1)}{\sigma_{f_{i,j}^0}} U\left(f_{i,j}^0\right)\right]^2 \quad (4.9)$$

$$\sigma_{i,SA}{}^2 = \left(\sum_{j=1}^{J} \sigma_{f_{ij}} S_j^{hyb}\right)^2 \quad (4.10)$$

$$U\left(f_{ij}^0\right) = \begin{matrix} 1, & |x| > 0 \\ 0, & x = 0 \end{matrix} \quad (4.11)$$

$$\sum_{i=1}^{I} r_{i,j} f_{i,j}^0 \leq 1 \quad (4.12)$$

$\sigma_{i,SA}{}^2$ is the uncertainty of the source impacts on species *i* (Eq. 4.10)*;* σ_{fij} is the numerical uncertainty of the reference source profiles; α is a sensitivity term $\alpha = \frac{I}{\sum_{i=1}^{I} \sum_{j=1}^{J} U(f_{ij}^0)}$ in this study); Γ is a numerical weighting term ($\Gamma = 0.01$ in this study); and $U\left(f_{ij}^0\right)$ is a piecewise function used to constrain the optimization to omit species with zero contribution from the source from having numerical influence on the optimization (Eq. 4.11). After optimization of the source profile coefficients (r_{ij}), revised source profiles (f_{ij}^{new}) are derived using Eq. 4.6. The second term serves as the error balancing term.

The reference source profiles f_{ij}^0 are derived from the study by Reff *et. al* (2009) [63]. The f_{ij}^0 matrix is composed of source fingerprints for 19 well-measured species (EC, Na, Al, Si, Cl, K, Ca, Ti, Mn, Fe, Cu, Zn, As, Se, Br, Sn, Sb, Ba, and Pb) (Tables C.2 and C.3). Secondary species OC, NO_3, NH_4, and SO_4 are omitted from the optimization because the method is targeted for species that are mainly primary because these species are inert in the CMAQ model and do not undergo chemical transformation after being emitted. After optimization, reference fractions of OC, NO_3, NH_4, and SO_4 are reincorporated into the optimized source profile. The original ratios of these four species are maintained in order to generate source profiles that are representative of the

fraction of primary species emitted. The final optimized source profile contains fingerprints for 20 sources and 23 species, including the 19 primary species and 4 secondary species.

The source profile optimization methods are applied over all CSN sites with complete sets of speciated data (no missing species). After optimization, new source profiles are grouped according to season (winter, spring, summer, and fall) and their associated EPA administrative region (Fig. 4.1). Region 4 had the greatest number of monitors (N = 56), and Region 7 had the least number of monitors (N = 4) (Table C.1).

The CMB-iteration method is used to evaluate the optimized source profiles for two cities: Atlanta, GA (Region 4) and St. Louis MO (Region 7). Implementation of the updated source profiles in CMB-iteration demonstrates the ability to use CTM-based profiles in traditional receptor models. The CMB-iteration method is a modification of the EPA CMB method and estimates both primary and secondary organic carbon impacts. Sources are selected to better allow comparison with results from previous studies for Atlanta and St. Louis [36, 87].

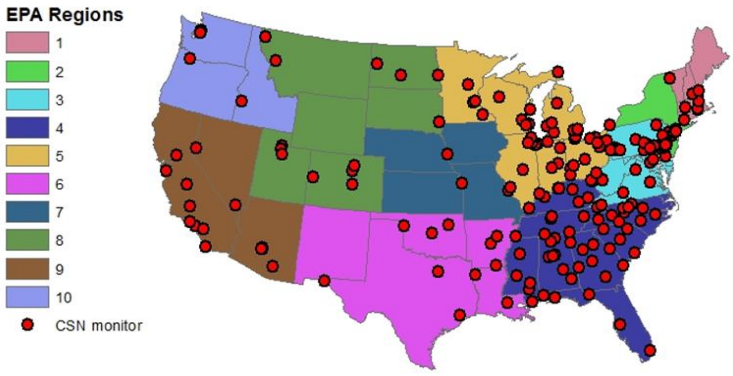

Figure 4.1: U.S. EPA administrative regions and CSN monitors used for model development and evaluation.

4.3 Results and Discussion

4.3.1 Source Profile Ratios

The ratios r_{ij} calculated as a result of the source profile optimization are presented in Figures C.1-C.20 in the form of seasonal distributions for each source and species. Distributions range from 0.5 to 2 as reflected in the constraints on the optimization. An r_{ij} value of 0.5 indicates a 50% decrease in the species contribution compared to the reference profile, and an r_{ij} of 2 indicates a doubling in species contribution to that source profile compared to the reference profile. Species fractions for agriculture/livestock (Fig. C.1) and biogenic (Fig. C.3) sources were less than 1 for EC, Na, Al, and Si (further designated at EC-Si), and were near one for all other species, indicating little variability in optimized ratios. Note that the reference contributions are zero for agriculture/livestock and biogenics, and contributions remain zero after optimization (no primary emissions of $PM_{2.5}$ from these sources).

Distributions for the other 18 sources had common characteristics. For example, ratios for EC-Si were less than 1 for all sources, indicating a reduction in these species' contributions after optimization. For fuel oil combustion (Fig. C.7) and non-road gasoline (Fig. C.12) ratios for the all other metals were mostly near one, indicating little change in species contribution for these sources. For coal combustion (Fig. C.4), dust (Fig. C.5), fires (Fig. C.6) metals processing (Fig. C.9), on-road gasoline combustion (Fig. C.15), other $PM_{2.5}$ (Fig. C.16), and other combustion (Fig. C.17), the distributions for most metals (Ca, Ti, Mn, Zn, Cu, Se, Br, and Sn) spanned the boundary constraints for r_{ij}, from 0.5 to 2. This indicates that metals contributions for these sources underwent more change than for the other sources. Each source had unique patterns of distributions for each species, indicating that the source profile optimization captures local variabilities that are unique to each source.

4.3.2 Optimized Source Profiles

A total of 1955 source profiles were calculated in the application of the source profile optimization method, and profiles were stratified by EPA region and season. All profiles were also averaged to calculate a national average profile. Contributions and standard deviations for the U.S. average for all sources are presented in the supplemental information (Tables C.3-C.10). The cosine similarity for source *j* was calculated for the optimized profile compared with the reference Reff *et al*. (2009) profile (Eq. 4.8) [63].

$$similarity_j = \frac{f_{ij}^{new} \cdot f_{ij}^{0}}{\left\| f_{ij}^{new} \right\| \left\| f_{ij}^{0} \right\|} = \frac{\sum_{i=1}^{I} f_{ij}^{new} f_{ij}^{0}}{\sqrt{\sum_{i=1}^{I} (f_{ij}^{new})^2} \sqrt{\sum_{i=1}^{I} (f_{ij}^{0})^2}} \quad (4.8)$$

A similarity of 1 indicates that the profiles are the same, -1 indicates opposite profiles, and 0 indicates orthogonality [88].

Profiles for sources and regions with the largest variability, or lowest similarity compared to original profiles, are highlighted (Tables C.11-C.13). Seasonal coal combustion profiles for Regions 3, 7, 10, and the entire U.S. are presented (Fig. 4.2). Similarities between Region 3 profiles and the original Reff *et al.* profile for coal combustion were 1.00 (winter), 0.97 (spring), 0.98 (summer), and 1.00 (fall), where the greatest variability occurred for warmer season profiles (Table C.11). The similarities for Region 7 coal combustion profiles were 0.98 (winter), 0.97 (spring), 0.98 (summer), and 0.99 (fall), where the greatest variability was also seen in the profiles for spring. The similarities for Region 10 coal combustion profiles were 0.99 (winter), 0.98 (spring), and 0.99 (summer), and the greatest variability was seen in the spring. Observations were not found during the fall months for Region 10 monitors, which led to an omission of this season/region from the source profile development. Of all regions, Region 7 profiles exhibited the greatest variability in seasonal profiles for coal combustion. Variability in coal combustion profiles is largely driven by bias in Fe concentrations.

The similarities between Region 3 profiles and the original Reff *et al.* profile for metals processing were 0.95 (winter), 0.97 (spring), 0.97 (summer), and 0.95 (fall), where the greatest variability occurred for the colder season profiles (Fig. 4.3, Table C.12). The similarities for Region 4 profiles were 0.96 (winter), 0.97 (spring), 0.97 (summer), and 0.95 (fall), where the greatest variability occurred for cooler season profiles. The similarities for Region 5 profiles were 0.96 (winter), 0.96 (spring), 0.93 (summer), and 0.93 (fall), where the greatest variability occurred for summer and fall profiles. Of all regions, profiles for Regions 5 exhibited the greatest variability in seasonal profiles for metals processing. Variability in metals processing profiles is largely driven by bias in Cl concentrations.

The similarities between Region 1 profiles and the original Reff *et al.* profile for sea salt were 0.75 (winter), 0.92 (spring), 0.77 (summer), and 0.74 (fall), where the greatest variability occurred for the fall profiles (Fig. 4.4). The similarities for Region 2 profiles were 0.61 (winter), 0.78 (spring), 0.61 (summer), and 0.67 (fall), where the greatest variability occurred for the winter and summer profiles. The similarities for Region 9 profiles were 0.99 (winter), 0.87 (spring), 0.52 (summer), and 0.67 (fall), where the greatest variability occurred for summer profiles. Of all regions, Region 9 profiles exhibited the greatest variability in seasonal profiles for sea salt. Variability in sea salt profiles is also largely driven by bias in Cl concentrations.

4.3.3 Revised Concentrations

Concentrations of metals were compared for observations, SH concentrations derived from the reference source profile ($SH_{reference}$), and SH concentrations derived from the revised source profiles ($SH_{revised}$) (Table C.14). Overall, correlations between observed and modeled concentrations of the most biased species improved, namely Cl (correlation: obs v. $SH_{reference}$ = 0.16, obs v. $SH_{revised}$ = 0.35; mean normalized bias (MNB): obs v. $SH_{reference}$ = 12.72, obs v. $SH_{revised}$ = 4.10) (Table C.15). Figures C.21 and

C.22 show a tighter spread for revised biases compared to reference biases. Some species simulation performance did not improve as significantly as Cl due to the overwhelmingly large bias in Cl compared to concentrations. However improvements were seen for Al, Si, K, Ca, Ti, and Se. Overall the revised profiles did lead to improved estimates of metals concentrations.

4.3.4 Receptor Model Application

The CMB-iteration method, which is a modification of the original CMB method by the EPA that estimates both primary and secondary organic carbon contributions, was applied for Atlanta, GA and St. Louis, MO for January and July 2006 to determine the performance of the optimized profiles in traditional receptor models (Fig. 4.5). The Region 4 winter and summer profiles were used for Atlanta, and the Region 7 winter and summer profiles were used for St. Louis. Monthly-averaged CMB-iteration results were compared with results from CMB-gas constraints (CMB-GC) applications and showed similar results [36]. Source impacts are also presented for the application of the spatial hybrid method using the reference ($SH_{reference}$) and revised ($SH_{revised}$) sources profiles. In general, the sources impacts from CMB-GC and CMB-iteration had similar trends while the $SH_{reference}$ and $SH_{revised}$ impacts had similar trends; however impacts for certain sources and seasons for all four methods were similar.

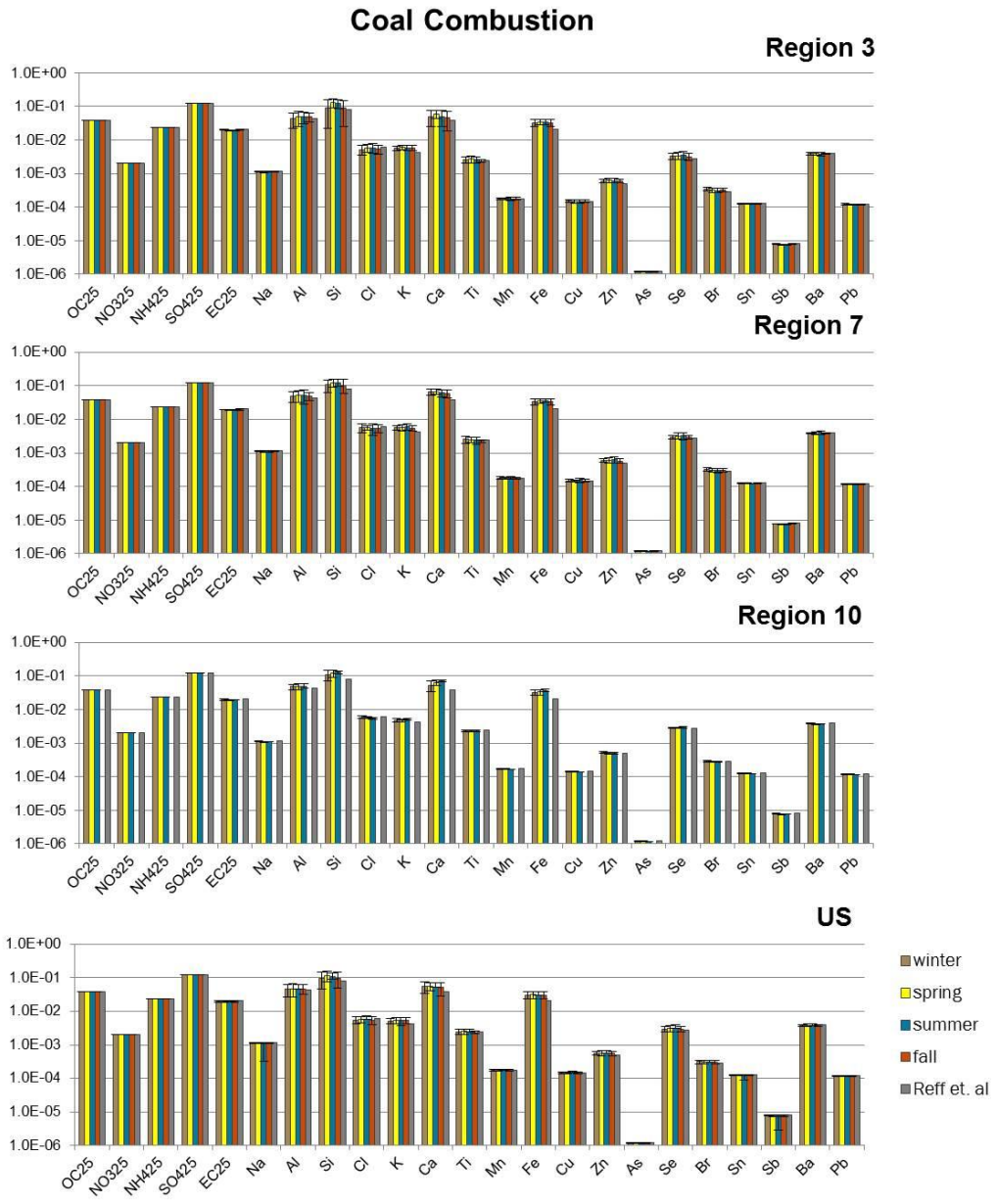

Figure 4.2: Coal combustion profiles (in mass fractions) for $PM_{2.5}$ for EPA Regions 3, 7, and 10 and the entire U.S. No data was available in fall for Region 10 monitors.

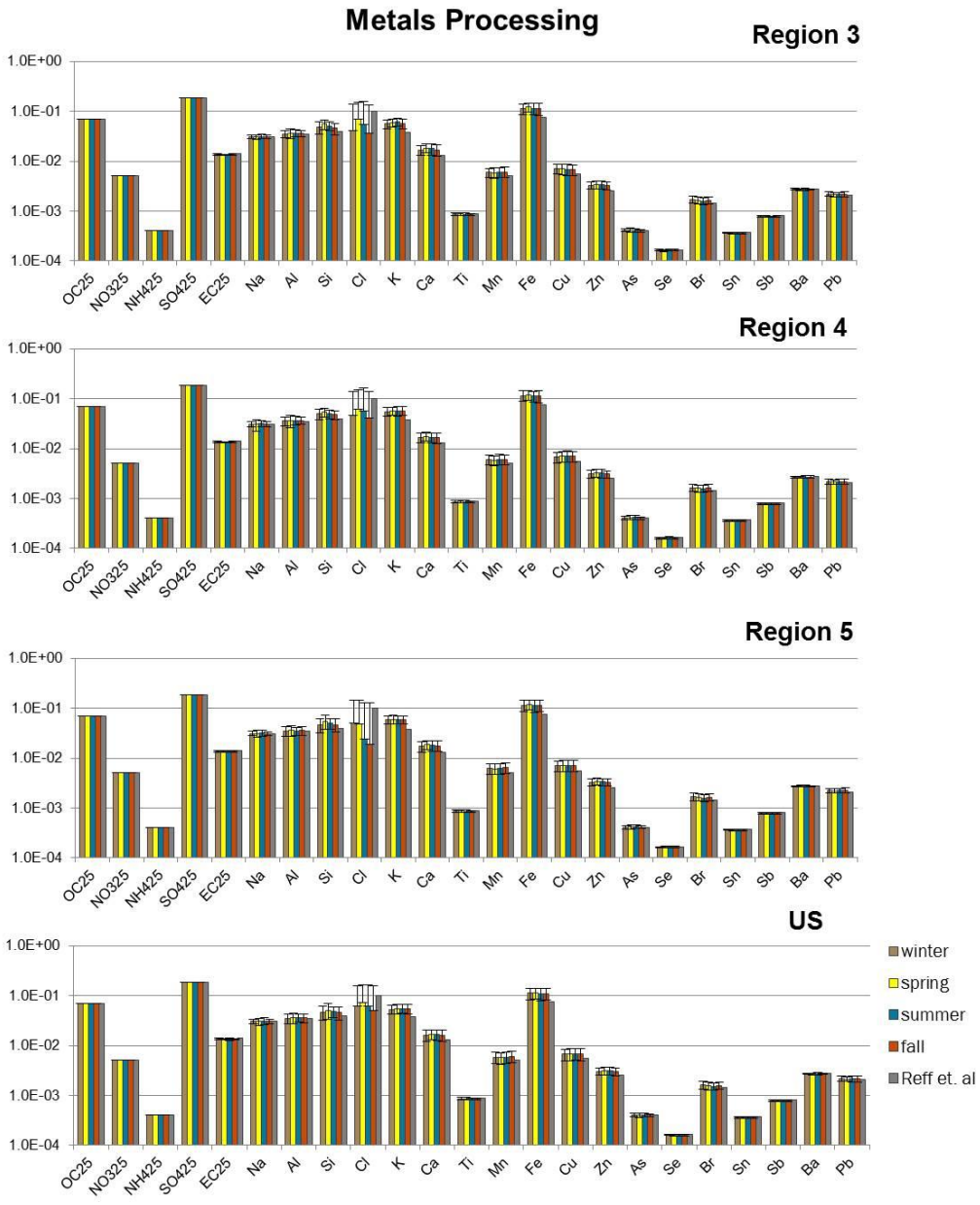

Figure 4.3: Metals processing profiles (in mass fractions) for $PM_{2.5}$ for EPA Regions 3, 4, and 5 and the entire U.S.

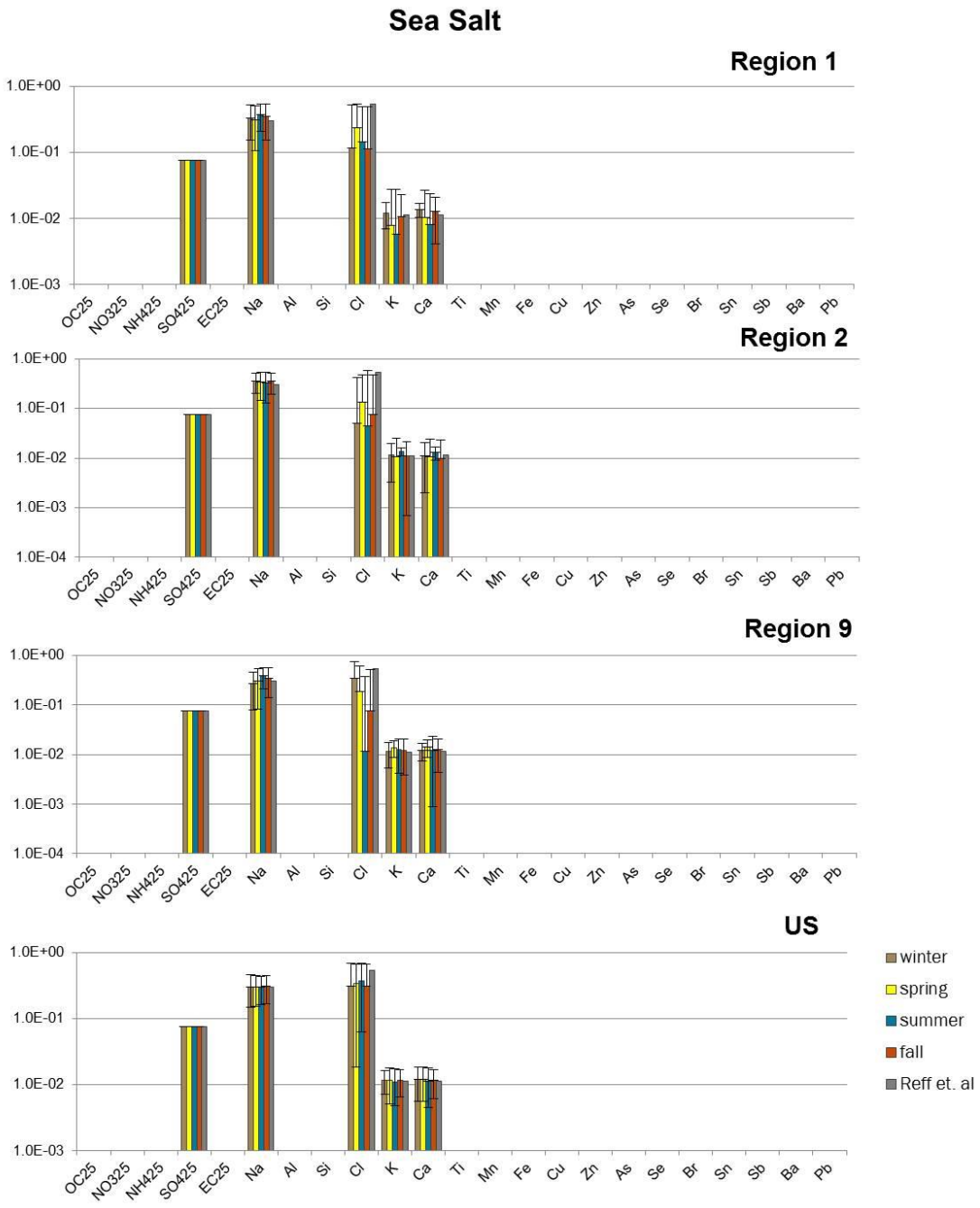

Figure 4.4: Sea salt profiles (in mass fraction) for $PM_{2.5}$ for EPA Regions 5, 6, and 9 and the entire U.S.

For Atlanta dust impacts, CMB-iteration had the highest impacts followed by CMB-GC, and the impacts from the SH applications were significantly lower than the impacts from the CMB applications (CMB-GC, winter: 0.14 ± 0.05 $\mu g \cdot m^{-3}$, summer: 1.25 ± 0.11$\mu g \cdot m^{-3}$; CMB-iteration, winter: 1.08 ± 0.51 $\mu g \cdot m^{-3}$, summer: 2.41 ± 0.32 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 0.07 ± 0.09 $\mu g \cdot m^{-3}$, summer: 0.14 ± 0.26 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 0.07 ± 0.04 $\mu g \cdot m^{-3}$, summer: 0.15 ± 0.28 $\mu g \cdot m^{-3}$). For all methods, dust impacts were higher during the summer season. For biomass burning impacts, the CMB impacts were again higher than SH impacts (CMB-GC, winter: 1.11 ± 0.54 $\mu g \cdot m^{-3}$, summer: 1.81 ± 0.08 $\mu g \cdot m^{-3}$; CMB-iteration, winter: 1.43 ± 0.30 $\mu g \cdot m^{-3}$, summer: 0.76 ± 1.08 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 0.72 ± 0.35 $\mu g \cdot m^{-3}$, summer: 0.02 ± 0.04 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 0.75 ± 0.35 $\mu g \cdot m^{-3}$, summer: 0.02 ± 0.04 $\mu g \cdot m^{-3}$). Summer time burning impacts from SH applications were especially low compared to CMB impacts. The discrepancy in dust and burn impacts between CMB and SH applications could be attributed by the higher number of sources analyzed in the SH applications, leading to more primary mass being apportioned to the "others" category (See Fig 4.5). Also, the SH methods tend to reduce the impacts from biomass burning and dust during SH optimization due to the high positive bias in tracers species concentrations for the sources (e.g., K for biomass burning and Si for dust).

For Atlanta coal combustion impacts, impacts from SH methods were much higher than impacts from CMB methods (CMB-GC, winter: 0.01 ± 0.04 $\mu g \cdot m^{-3}$, summer: 0.11 ± 0.09 $\mu g \cdot m^{-3}$; CMB-iteration, winter: -1.27 ± 0.84 $\mu g \cdot m^{-3}$, summer: -1.20 ± 0.84 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 0.32 ± 0.18 $\mu g \cdot m^{-3}$, summer: 0.58 ± 0.18 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 0.35 ± 0.19 $\mu g \cdot m^{-3}$, summer: 0.61 ± 0.30 $\mu g \cdot m^{-3}$). For SH impacts, summertime impacts were higher than winter impacts. The CMB-iteration method did not resolve the coal combustion source impacts, leading to negative contributions on most days. The source-oriented approach is better able to capture the combined primary and secondary

impacts from coal combustion compared to the receptor oriented approaches that separate primary and secondary impacts.

For the acidic secondary sources, results are comparable across the methods: ammonium sulfate (CMB-GC, winter: 1.37 ± 0.17 $\mu g \cdot m^{-3}$, summer: 5.32 ± 0.45 $\mu g \cdot m^{-3}$; CMB-iteration, winter: 1.37 ± 0.62 $\mu g \cdot m^{-3}$, summer: 5.44± 0.62 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 2.03 ± 0.99 $\mu g \cdot m^{-3}$, summer: 2.89 ± 1.12 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 2.03 ± 0.99 $\mu g \cdot m^{-3}$, summer: 2.89 ± 1.12 $\mu g \cdot m^{-3}$), ammonium bisulfate (CMB-GC, winter: 1.12 ± 0.19 $\mu g \cdot m^{-3}$, summer: 2.50 ± 0.52 $\mu g \cdot m^{-3}$; CMB-iteration, winter: 1.41 ± 0.67 $\mu g \cdot m^{-3}$, summer: 2.60 ± 0.67 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 2.18 ± 1.06 $\mu g \cdot m^{-3}$, summer: 3.13 ± 1.22 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 2.18 ± 1.06 $\mu g \cdot m^{-3}$, summer: 3.13 ± 1.22 $\mu g \cdot m^{-3}$), and ammonium nitrate (CMB-GC, winter: 1.02 ± 0.09 $\mu g \cdot m^{-3}$, summer: 0.42 ± 0.04 $\mu g \cdot m^{-3}$; CMB-iteration, winter: 1.01 ± 0.10 $\mu g \cdot m^{-3}$, summer: 0.42 ± 0.10 $\mu g \cdot m^{-3}$; $SH_{reference}$, winter: 0.78 ± 0.64 $\mu g \cdot m^{-3}$, summer: 0.48 ± 0.31 $\mu g \cdot m^{-3}$; $SH_{revised}$, winter: 0.78 ± 0.64 $\mu g \cdot m^{-3}$, summer: 0.48 ± 0.31 $\mu g \cdot m^{-3}$). The SH methods have slightly lower ammonium sulfate impacts and slightly higher ammonium bisulfate impacts compared to CMB methods because they were derived directly from sulfate, nitrate, and ammonium concentrations for comparison. Also, the acid impacts for SH methods are the same because secondary species concentrations are unchanged in the source profile optimization. For all methods, sulfate impacts are higher than nitrate impacts.

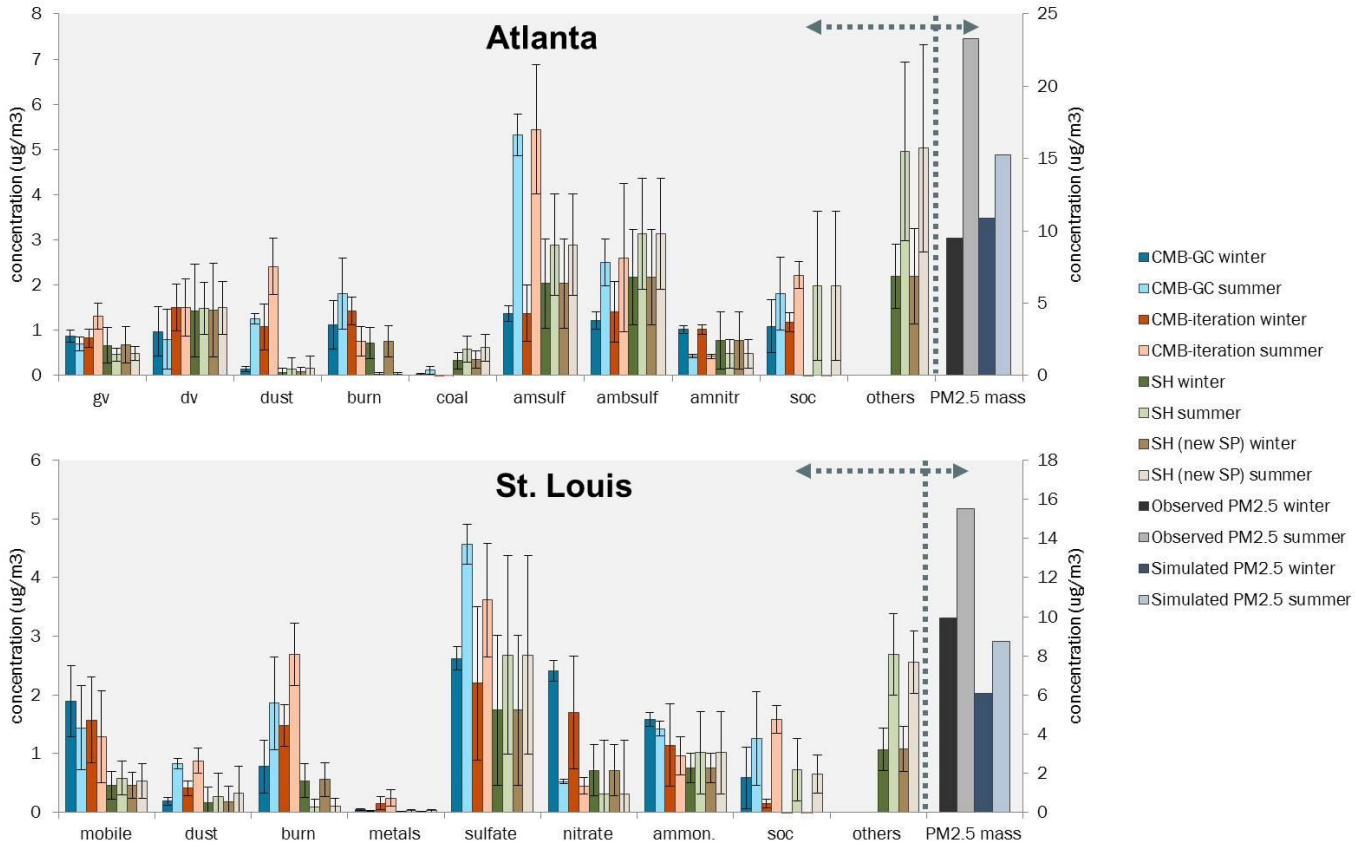

Figure 4.5: Averaged CMB-GC, CMB-iteration, original spatial hybrid (SH), and spatial hybrid with new source profiles (new SP) results for a winter and summer month for Atlanta, GA and St. Louis, MO. CMB-GC results are obtained from an application of methods from Balachandran et al. (2012) and Maier et al. (2013) [36, 87]. $PM_{2.5}$ mass concentration data is obtained from the CSN network. The arrows indicate the direction of the appropriate y-axis for the bars on either side of the dashed line

For Atlanta SOC, results are similar over all methods with the exception of winter time SOC for SH methods (CMB-GC, winter: 1.09 ± 0.53 μg·m^{-3}, summer: 1.81 ± 0.81 μg·m^{-3}; CMB-iteration, winter: 1.18 ± 0.21 μg·m^{-3}, summer: 2.22 ± 0.21 μg·m^{-3}; $SH_{reference}$, winter: -0.83 ± 0.51 μg·m^{-3}, summer: 1.98 ± 1.65 μg·m^{-3}; $SH_{revised}$, winter: -0.83 ± 0.51μg·m^{-3}, summer: 1.98 ± 1.65 μg·m^{-3}). The winter SOC impacts for SH methods were negative due to low bias in CMAQ-modeled SOC in winter for this study, which does not impact the receptor-oriented methods. The modeled $PM_{2.5}$ for Atlanta was biased low in summer, which may be attributed to the low bias in dust and biomass burning impacts caused by biases further upstream in the flow of the presented methods.

St. Louis, MO

For St. Louis, MO, the same sources as studied by Maier et al. (2013) were chosen for analysis in CMB-iteration: mobile (sum of gasoline and diesel impacts), dust, biomass burning (burn), metals, sulfate, nitrate, ammonium, and SOC (Fig. 4.5) [87]. The Maier et al. (2013) study included metal processing as a source for St. Louis due to the large presence of metal-working industries in the area [87]. Impacts from CMB methods had similar trends and impacts from SH methods had similar trends. Scatter plots for each source category and method for winter and summer seasons are presented in Figures C.25 and C.26, respectively. Overall, the SH impacts were lower than the CMB impacts because the CSN site used in the SH application is located farther from the city center than the Blair Street monitor used in the comparison study. The seasonality of the mobile source impacts was opposite for CMB and SH methods in that CMB impacts are higher in winter and SH impacts are higher in summer (CMB-GC, winter: 1.89 ± 0.61 μg·m^{-3}, summer: 1.44 ± 0.72 μg·m^{-3}; CMB-iteration, winter: 1.57 ± 0.73 μg·m^{-3}, summer: 1.29 ± 0.78 μg·m^{-3}; $SH_{reference}$, winter: 0.46 ± 0.24 μg·m^{-3}, summer: 0.58 ± 0.29 μg·m^{-3}; $SH_{revised}$, winter: 0.46 ± 0.22 μg·m^{-3}, summer: 0.53 ± 0.30 μg·m^{-3}).

For St. Louis dust impacts, impacts are higher for CMB methods and are higher in the summer for all four methods (CMB-GC, winter: 0.19 ± 0.07 μg·m^{-3}, summer: 0.83 ± 0.09 μg·m^{-3}; CMB-iteration, winter: 0.41 ± 0.12 μg·m^{-3}, summer: 0.88 ± 0.21 μg·m^{-3}; $SH_{reference}$, winter: 0.17 ± 0.26 μg·m^{-3}, summer: 0.27 ± 0.40 μg·m^{-3}; $SH_{revised}$, winter: 0.18 ± 0.26 μg·m^{-3}, summer: 0.33 ± 0.45 μg·m^{-3}). For biomass burning impacts, CMB impacts are significantly higher than SH impacts, and CMB-iteration impacts are highest (CMB-GC, winter: 0.78 ± 0.45 μg·m^{-3}, summer: 1.86 ± 0.79 μg·m^{-3}; CMB-iteration, winter: 1.48 ± 0.36 μg·m^{-3}, summer: 2.70 ± 0.53 μg·m^{-3}; $SH_{reference}$, winter: 0.54 ± 0.29 μg·m^{-3}, summer: 0.09 ± 0.13 μg·m^{-3}; $SH_{revised}$, winter: 0.56 ± 0.28 μg·m^{-3}, summer: 0.10 ± 0.14 μg·m^{-3}). Summer time burning impacts are higher for CMB methods and winter time burning impacts are higher for SH methods. Similar to Atlanta, dust and biomass burning impacts are biased low due to a low bias in emissions or a high bias in modeled tracer species concentrations. Also, primary mass could be apportioned to the "others" category for the SH methods. Metals impacts were similar for all methods but slightly higher for CMB-iteration (CMB-GC, winter: 0.05 ± 0.01 μg·m^{-3}, summer: 0.03 ± 0.01 μg·m^{-3}; CMB-iteration, winter: 0.16 ± 0.11 μg·m^{-3}, summer: 0.24 ± 0.14 μg·m^{-3}; $SH_{reference}$, winter: 0.01 ± 0.01 μg·m^{-3}, summer: 0.02 ± 0.02 μg·m^{-3}; $SH_{revised}$, winter: 0.01 ± 0.01 μg·m^{-3}, summer: 0.02 ± 0.02 μg·m^{-3}).

For impacts from acidic secondary sources, CMB impacts were higher than SH impacts: sulfate (CMB-GC, winter: 2.62 ± 0.19 μg·m^{-3}, summer: 4.57 ± 0.33 μg·m^{-3}; CMB-iteration, winter: 2.22 ± 1.31 μg·m^{-3}, summer: 3.62 ± 0.97 μg·m^{-3}; $SH_{reference}$, winter: 1.74 ± 1.28 μg·m^{-3}, summer: 2.68 ± 1.69 μg·m^{-3}; $SH_{revised}$, winter: 1.74 ± 1.28 μg·m^{-3}, summer: 2.68 ± 1.69 μg·m^{-3}), nitrate (CMB-GC, winter: 2.41 ± 0.17 μg·m^{-3}, summer: 0.52 ± 0.04 μg·m^{-3}; CMB-iteration, winter: 1.70 ± 0.96 μg·m^{-3}, summer: 0.45 ± 0.13 μg·m^{-3}; $SH_{reference}$, winter: 0.71 ± 0.44 μg·m^{-3}, summer: 0.32 ± 0.92 μg·m^{-3}; $SH_{revised}$, winter: 0.71 ± 0.44 μg·m^{-3}, summer: 0.32 ± 0.92 μg·m^{-3}), and ammonium (CMB-GC,

winter: 1.58 ± 0.12 μg·m^{-3}, summer: 1.43 ± 0.12 μg·m^{-3}; CMB-iteration, winter: 1.15 ± 0.71 μg·m^{-3}, summer: 0.96 ± 0.33 μg·m^{-3}; $SH_{reference}$, winter: 0.76 ± 0.25 μg·m^{-3}, summer: 1.02 ± 0.70 μg·m^{-3}; $SH_{revised}$, winter: 0.76 ± 0.25 μg·m^{-3}, summer: 1.02 ± 0.70 μg·m^{-3}). Seasonality was similar for all methods for sulfate and nitrate; however CMB ammonium impacts were higher in winter and SH ammonium impacts were higher in summer.

For SOC, impacts were similar for all methods and higher in the summer (CMB-GC, winter: 0.59 ± 0.53 μg·m^{-3}, summer: 1.26 ± 0.80 μg·m^{-3}; CMB-iteration, winter: 0.15 ± 0.07 μg·m^{-3}, summer: 1.58 ± 0.24 μg·m^{-3}; $SH_{reference}$, winter: -0.32 ± 0.24 μg·m^{-3}, summer: 0.73 ± 0.53 μg·m^{-3}; $SH_{revised}$, winter: -0.32 ± 0.24 μg·m^{-3}, summer: 0.73 ± 0.53 μg·m^{-3}). SOC impacts were negative in winter for SH methods due to low bias in CMAQ-modeled SOC in winter in this study.

4.3.5 Implications

This analysis demonstrates that CTM-RM-derived source profiles can be used in traditional receptor model studies. However, the traditional methods are unable to make use of all of the source categories at once due to source profile similarities, which leads to negative source contributions, due, in part, to overfitting by the model. It is suggested that the most appropriate source categories be chosen when performing traditional receptor modelling with optimized source profiles. Additionally, optimized source profiles can be utilized in chemical transport models by providing transport based information into the calculation of primary trace species concentrations. By optimizing source profiles over several monitors, the understanding of the uncertainty in each profile increases. This study produced fairly consistent source profiles over the entire U.S., where concentrations and some source impacts have significant spatial and temporal variability. Results also imply that spatial variability in source profiles is an important factor to consider when choosing source profiles. Improvement in modeled trace metal concentrations was also seen in comparison to observations.

4.4 Summary

A nonlinear optimization method for calculating new $PM_{2.5}$ source profiles is developed and applied over all available CSN monitors for the year 2006. The method takes into account both observations and modeled source impact estimates to produce receptor-based source profiles for 20 source categories and 23 $PM_{2.5}$ component species. The results indicate that there is both seasonal and spatial variation in optimized source profiles. An application of the source profiles in a traditional receptor model produced more informative results in comparison to other peer-reviewed receptor modeling studies. The updated source profiles can provide detailed location-based species fractions that are beneficial for determining receptor-based impacts of trace metals from important $PM_{2.5}$ sources.

The CMB-iteration and source profile optimization methods have been formatted for wide-scale use in the form of a user-friendly GUI (Figs. C.27 and C.28). The programs (CMB-iteration 3.0 and SSAPO (Simultaneous Source Apportionment with Profile Optimization)) can be used to apply the methods for one location and one time. Both programs are available for download from Georgia Tech's Russell group website (http://russellgroup.ce.gatech.edu/node/16). The source profile optimization method is applicable for any location where CTM sensitivities, speciated observations, and reference profiles are available. The reference profile serves as an initial guess for the calculation of a locally-based source profile. While locally-based profiles are traditionally difficult to obtain, especially for developing regions such as China and India, observed and modeled data are relatively easier to obtain. New optimized source profiles may better reflect local emissions sources and pollutant characteristics.

4.5 Acknowledgements

This publication was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901, STAR Fellowship FP-91761401-0, and by NASA under grant NNX11AI55G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Alfred P. Sloan Foundation for their support.

CHAPTER 5

A METHOD FOR QUANTIFYING BIAS IN MODELED CONCENTRATIONS AND SOURCE IMPACTS FOR SECONDARY PARTICULATE MATTER

As published in *Frontiers in Environmental Science and Engineering*

Abstract

CMAQ estimates of sulfates, nitrates, ammonium, and organic carbon are highly influenced by uncertainties in modeled secondary formation processes, such as chemical mechanisms, volatilization, and condensation rates. These compounds constitute the majority of $PM_{2.5}$ mass, and reducing bias in estimated concentrations has benefits for policy measures and epidemiological studies. In this work, a method for adjusting source impacts on secondary species is developed that provides estimates of source contributions and reduces bias in modeled concentrations compared to observations. The bias correction adjusts concentrations and source impacts based on the difference between modeled concentrations and observations while taking into account uncertainties at the location of interest; and it is applied both spatially and temporally. We apply the method over the U.S. for 2006. The mean bias for initial CMAQ concentrations compared to observations is -0.28 (OC), 0.11 (NO_3), 0.05 (NH_4), and -0.08 (SO_4). The normalized mean bias in modeled concentrations compared to observations was effectively zero for OC, NO3, NH4, and SO4 after applying the secondary bias correction. Ten-fold cross-validation was conducted to determine the performance of the spatial application of the bias correction. Cross-validation performance was favorable; correlation coefficients were greater than 0.69 for all species when comparing observations and concentrations

based on kriged correction factors. The methods presented here address model uncertainties by improving simulated concentrations and source impacts of secondary particulate matter through data assimilation. Secondary-adjusted concentrations and source impacts from 20 emissions sources are generated for 2006 over continental U.S.

5.1 Introduction

Ambient concentrations of particulate matter with an aerodynamic diameter of 2.5 microns or less ($PM_{2.5}$) are regulated in the United States by the National Ambient Air Quality Standards, which are set by the U.S. Environmental Protection Agency. Ambient $PM_{2.5}$ can be composed of hundreds of different chemical species which originate from both natural and anthropogenic sources. The majority of $PM_{2.5}$ mass is composed of sulfates, ammonium, nitrates, and organic carbon (SANOC), as well as water. The remaining mass is composed of sodium, chloride, and other minor inorganic species, including trace metals and their oxides. SANOC is largely formed in the atmosphere from reactions involving precursor species, though a fraction can be directly emitted. Modeling ambient $PM_{2.5}$ concentrations, particularly SANOC, is challenging due to difficulties in capturing complex atmospheric interactions [89-95].

Several studies evaluated the performance of modeling ambient SANOC concentrations. Pandis *et al*. (1992) evaluated the application of a Lagrangian model for estimating secondary organic aerosol (SOA) concentrations in the Los Angeles area [96, 97]. The model estimated the percentage of SOA yield due to photo-oxidation of aromatics (65%), biogenic hydrocarbons (16%), alkanes (15%), and olefins (4%). The study also examined model responses to deposition velocity, emissions of reactive organic gases (ROGs) and NOx, and saturation concentrations. In the Pandis *et al.* study, authors reported that the major modeling uncertainties originated from ROG emissions, aerosol yields, and the partitioning of condensable gases between the gas and aerosol phase. Hildemann *et al.* (1993) also evaluated the modeling performance of the

Lagrangian model in estimating SOA formation in the Los Angeles basin by comparing modeled estimates to high-resolution gas chromatography observations of organic aerosol concentrations [98]. Results indicated that the model underpredicted acidic organic aerosol concentrations due to the absence of secondary processes from the model or missing emissions of precursors in the inventory used.

Simon *et al.* (2009) presented a comprehensive review of 69 operational evaluation studies published from 2006 to 2012 that evaluated photochemical model performance in simulating concentrations and wet deposition of sulfate, nitrate, and ammonium (SNA), as well as the performance in simulating ozone and fine and coarse PM concentrations [99]. Authors reported that model performance for simulating SNA in the eastern U.S. is better than for simulations in the western U.S. Authors attributed the regional discrepancies to the predominance of model evaluations over eastern regions, which resulted in the improvement of model inputs and processes for the region. Measured and simulated sulfate was generally lower in the western U.S., also attributing to higher modeled biases in the west. Further, complex meteorology and terrain in the west contributed to difficulties in modeling high nitrate and organic carbon episodes.

Kanakidou *et al*. (2005) presented a comprehensive review which investigated important questions regarding organic aerosol modeling and its relationship to global climate [100]. The authors detailed uncertainties in biogenic emissions of organic aerosols due to land cover changes over time and variability in plant species' emissions patterns. Emissions of anthropogenic organics were uncertain due to uncertainty in activity profiles of fuel use and future emission factors. The review discussed the uncertainties in the knowledge, at that time, of the atmospheric chemistry of organics, including gas-phase photo-oxidation processes, semi-volatile aerosol and oligomer formation, and the impact of NO_x on secondary organic aerosol formation. Finally, the review discussed uncertainties in the aerosol physics processes, such as gas-aerosol partitioning and the role of organic aerosols in cloud microphysics and rain processes.

Bessagnet *et. al* evaluated the performance of the aerosol module in CHIMERE, a chemical transport model, and compared simulated SNA concentrations to observations over Europe [101]. The authors attributed a positive bias in sulfate concentrations to inflexible equations for aqueous chemistry, a positive bias in nitrate to over-evaporation during summer, and a positive bias in ammonium to deficiencies in horizontal transport and overestimated boundary conditions.

Overall, these studies found that modeling the secondary formation of atmospheric aerosols has many uncertainties. Therefore, bias correction methods are beneficial for improving the predicted secondary aerosol concentrations. This manuscript introduces a bias correction method for source impacts on SANOC, which in turn improves modeled concentration estimates. With the many sources of uncertainty involved in modeling SANOC, the methods presented are designed to provide a mathematical, physically relevant adjustment that integrates model and measurement uncertainties, modeled pollutant sensitivities, and modeled and observed concentrations in a way that closes the gap between modeled estimates and observations. While traditional receptor models are challenged to identify the sources of secondary $PM_{2.5}$ beyond gas-to-particle interactions, the sources of secondary $PM_{2.5}$ and the magnitude of the impacts are directly identified in this study.

5.2 Data and Methods

5.2.1 Data

This study uses speciated $PM_{2.5}$ observations from the Chemical Speciation Network (CSN) for model development and evaluation. Data are available every third or sixth day from CSN observation sites across the U.S. Total $PM_{2.5}$ mass and 40 species concentrations are used in this study for spatial hybrid source apportionment including SANOC, elemental carbon, and 35 trace metals. Days and sites with missing species

concentrations were discarded from the spatial hybrid application. CSN observations are available on 121 days in 2006 (maximum possible temporal availability for the year), and the number of monitors with a complete set of speciated data available on observation days varied from approximately 20 to 120 monitors. Carbon data from CSN was adjusted from TOT to TOR equivalents [75].

5.2.2 CMAQ–DDM Modeling

The Community Multi-Scale Air Quality (CMAQ) model is a chemical transport model that is equipped with the ability to estimate model sensitivities to perturbations in inputs or boundary conditions by implementing the decoupled direct method (DDM) for three-dimensional grids [46, 49, 76]. CMAQ-DDM modeling was performed for one year (2006) at 36-km resolution over continental U.S. and parts of southern Canada and northern Mexico (Fig. 5.1). Meteorological inputs were generated using the Weather Research and Forecasting (WRF) model v3.3.1, along with the Pleim-Xiu land-surface model, Kain-Fritsch cumulus parameterization, Morrison 2-moment micro physics, RRTM longwave radiation, Dudhia shortwave radiation, and the ACM2 planetary boundary layer scheme [60, 61]. Emissions inputs were generated using the Sparse Matrix Operator Kernel Emissions (SMOKE) model (v2.6) with the 2005 U.S. EPA National Emissions Inventory data grown to 2006 [77]. The CMAQ-DDM model was used to generate $PM_{2.5}$ sensitivities to 20 unique sources, and these sensitivities represent the initial source impact estimates (base case) [21, 52].

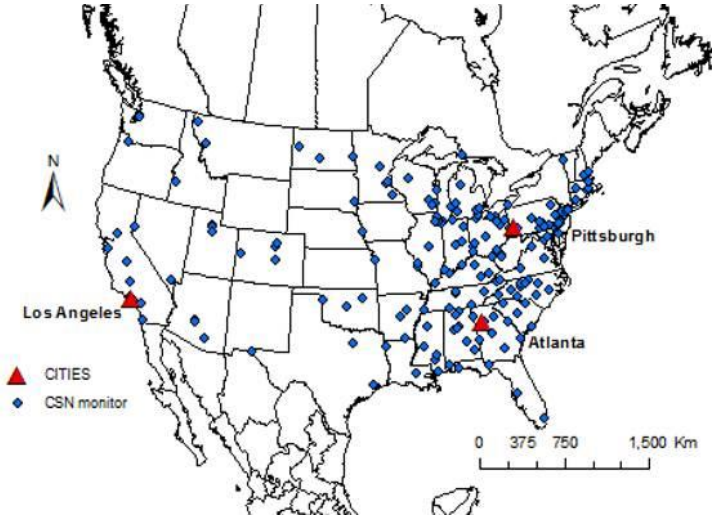

Figure 5.1: CSN monitors (blue circles) used for model development, application, and evaluation.

5.2.3 Spatial Hybrid Source Apportionment

The spatial hybrid (SH) method for spatial source apportionment combines geostatistical methods with sparsely located source impact estimates that have been corrected by observations using the hybrid, chemical transport model-receptor model (CTM-RM) model. The SH model generates spatial fields of concentrations and source impacts for total $PM_{2.5}$ and 40 $PM_{2.5}$ components and is summarized here [21]. The CTM-RM model (Eq. 5.1) is solved using nonlinear optimization, where the optimized parameter R_j can be interpreted as a factor (for each source, *j*) that is used to adjust the base case CMAQ-DDM source impact estimates.

$$X^2 = \sum_{i=1}^{N}\left[\frac{\left[\left(c_i^{obs} - c_i^{sim} - \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1)\right)\right]^2}{{\sigma_{i,obs}}^2 + {\sigma_{i,SP}}^2}\right] + \Gamma \sum_{j=1}^{J} \frac{ln(R_j)^2}{{\sigma_{\ln(R_j)}}^2} \quad (5.1)$$

Eq. 5.1 is applied at each CSN monitor on days when speciated $PM_{2.5}$ measurements are available (typically every third day). In Eq. 5.1, X^2 is the error to be minimized; c_i^{obs} and c_i^{sim} are the observed and CMAQ-simulated concentrations of species *i*; $SA_{i,j}^{base}$ is the base CMAQ-DDM source impact of source *j* on species *i* before hybrid adjustment; R_j is

the adjustment factor for source j to be optimized; $\sigma_{i,obs}$, $\sigma_{i,CTM}$, and $\sigma_{\ln(R_j)}$ are the uncertainties in observed concentrations, CMAQ-simulated concentrations, and source strength, respectively; and Γ weights the source strength and serves as a numerical balancing parameter.

Kriging is used to spatially interpolate hybrid adjustment factors (R_j) for each source category on days with available speciated $PM_{2.5}$ measurements. Due to the limited temporal availability of observation data, hybrid adjustment factors were temporally interpolated using grid-by-grid linear interpolation to allow for daily adjustment of source impact spatial fields, as opposed to adjusting fields solely on days with available data. After grid-by-grid linear interpolation for each source category, daily spatial fields of adjustment factors are applied to base CMAQ-DDM spatial fields, resulting in daily, optimized $PM_{2.5}$ source impact fields over continental U.S.

5.2.4 Secondary Bias Correction

SH-adjusted estimates of source impacts on SANOC are refined using a secondary species correction that reduces the modeled bias in the secondary formation of the species. The secondary correction is only applied to SANOC species (sulfate, ammonium, nitrate, and organic carbon). In this work, the adjustment for organic carbon (OC) is considered to be for the secondary portion. In other words, the initial modeled OC concentrations are composed of primary and secondary components, and the bias correction is assumed to account for additional/reduced secondary OC formation. The adjustment $SC_{i,j}$ represents the mass adjustment in $\mu g \cdot m^{-3}$ for the impact of source j on species i, and is calculated at the monitor as follows:

$$c_i^{hyb} = c_i^{sim} + \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1) \quad (5.2)$$

$$D_i = c_i^{obs} - c_i^{hyb} \quad (5.3)$$

$$SC_{i,j} = \frac{SA_{i,j}R_j}{\sum_{j=1}^{J} SA_{i,j}{}^{+}R_j} * D_i \text{ when } SA_{i,j} > 0 \text{ ; and } SC_{i,j} = 0 \text{ when } SA_{i,j} \leq 0 \quad (5.4)$$

As a note, the occurrence of $SA_{i,j} \leq 0$ was 62% for organic carbon, 47% for nitrate, 57% for ammonium, and 56% for sulfate. In Eq. 5.2, $SA_{i,j}$ are CMAQ-DDM source impacts (+ indicates positive impacts only), c_i^{obs} is the observed concentration, and c_i^{hyb} is the reconstructed hybrid concentration (Eq. 5.4). In Eq. 5.4, the difference (*D*) between the hybrid-adjusted concentrations and observations is weighted by the magnitude of the source impacts and then redistributed across the impacts. A unique $SC_{i,j}$ value is calculated for each species/source combination (4 secondary species and 20 source categories) on each observation day. The final adjusted concentrations and source impacts are estimated using Eqs. 5.5 and 5.6:

$$c_i^{new} = c_i^{sim} - \sum_{j=1}^{J} SA_{i,j}^{base} + \sum_{j=1}^{J} SA_{i,j}^{base} R_j + SC_{ij} \quad (5.5)$$

$$SA_{i,j}^{new} = SA_{i,j}^{base} R_j + SC_{ij} \quad (5.6)$$

Kriging is then used to spatially interpolate the SC_{ij} values calculated at the monitors over the spatial domain on observation days.

With this approach, only positively impacting sources on species concentrations are corrected, which makes the adjustment physically realistic as some CMAQ-DDM source impacts are negative. As an example, if coal combustion contributed 20% of the sulfate mass in CMAQ-DDM, the bias correction will apportion 20% of the sulfate residual to the secondary coal combustion impact (Eq. 5.6). The presented approach also assumes that the majority of the SNA is secondary, eliminating the need to split the primary and secondary portions for these species. The adjustment SC_{ij} can be viewed as a correction for source impacts due to uncertainties in secondary processes, such as oxidation rates, aerosol yields and volatility, as well as emissions and meteorology, though the latter two are addressed, in part, in the original SH method. In the preceding study (Ivey *et. al*, 2015), the adjustment factor R_j was developed as a correction for emissions estimates and meteorological factors to reduce the bias in predicted primary species concentrations [21].

The spatial extension of the secondary bias correction was evaluated using 10-fold cross-validation. In this cross-validation, all available CSN site-days (7955 in total) were randomly arranged into 10 groups, where 9 groups had 795 site-days and 1 group had 800. The kriging step was performed 10 times, and each time one group was removed from the sample used for kriging. Data from withheld site-days were compared to the corresponding kriged values, leading to 7955 evaluation pairs. Some observation days (days where 1-in-6 day observations are not available) had too few monitors after removing the selected monitors, and the kriging algorithm did not converge.

5.3 Results

5.3.1 Base Model Evaluation

Base CMAQ-DDM results of SANOC and $PM_{2.5}$ concentrations are compared to observations from the CSN network to evaluate base model performance (Table D2). Mean observed concentrations were 12.6 ± 7.63, 2.22 ± 1.76, 1.58 ± 2.29, 1.40 ± 1.18, and 3.13 ± 2.61 $\mu g \cdot m^{-3}$ for total $PM_{2.5}$, organic carbon, nitrate, ammonium, and sulfate, respectively. Mean CMAQ-DDM concentrations were 12.6 ± 8.58, 1.60 ± 1.27, 1.76 ± 2.55, 1.47 ± 1.22, and 2.89 ± 2.30 $\mu g \cdot m^{-3}$ for total $PM_{2.5}$, organic carbon, nitrate, ammonium, and sulfate, respectively. Mean bias (MB) for total $PM_{2.5}$ (MB = -0.001) was lower than the mean bias for secondary species (MB_{OC} = -0.28, MB_{NO3} = 0.11, MB_{NH4} = 0.05, MB_{SO4} = -0.08), reflecting the uncertainty of model processes for secondary species. The MB values presented here indicate an overall overestimation of nitrate and ammonium concentrations by the base model application and an overall underestimation of organic carbon and sulfate concentrations. Modeled organic carbon concentrations had the largest bias compared to the observations. However, when comparing the normalized mean error (NME) for total $PM_{2.5}$ and secondary species, the NME is similar

($NME_{PM2.5}$ = 0.46, NME_{OC} = 0.53, NME_{NO3} = 0.79, NME_{NH4} = 0.49, NME_{SO4} = 0.38). Organic carbon and nitrate had the largest error compared to observations.

5.3.2 Secondary Bias Correction

Secondary bias correction factors (SC_{ij}) were calculated for each source and species combination (4 species x 20 sources) at all available CSN site-days in 2006 (121 total observation days). Mean SC_{ij} values and standard deviations are presented in Table D3, along with related statistics from the observations (Table D4). Conclusions can be drawn by examining the magnitude of the SC_{ij} values. A positive SC_{ij} value indicates that the initial simulated concentration was biased low compared to the observation, and vice versa for a negative SC_{ij} value. For secondary OC, larger SC_{ij} values were obtained for biogenics (average SC_{ij} over all days and sites, μ = 0.87 μg·m^{-3}), non-road gasoline (μ = 0.036 μg·m^{-3}), on-road gasoline (μ = 0.026 μg·m^{-3}), and solvents (μ = 0.19 μg·m^{-3}), indicating that these are the significant sources for secondary OC and that increasing the impact from these sources improves modeled concentrations. For nitrate, larger SC_{ij} values were obtained for agriculture/livestock activities (μ = 0.038 μg·m^{-3}), natural gas combustion (μ = 0.031 μg·m^{-3}), on-road diesel (μ = 0.012 μg·m^{-3}), and on-road gasoline (μ = 0.079 μg·m^{-3}), indicating a strong association between nitrate and these sources. For ammonium, agriculture/livestock activities (μ = 0.038 μg·m^{-3}), coal combustion (μ = 0.025 μg·m^{-3}), fuel oil combustion (μ = 0.020 μg·m^{-3}), and on-road gasoline (μ = 0.022 μg·m^{-3}) had larger SC_{ij} values. For sulfate, coal (μ = 0.27 μg·m^{-3}) and fuel oil (μ = 0.045 μg·m^{-3}) combustion, on-road diesel (μ = 0.023 μg·m^{-3}), and $PM_{2.5}$ others (μ = 0.041 μg·m^{-3}) had larger SC_{ij} values.

Distributions of SC_{ij} values provide information on how source impacts on secondary species are influenced by model bias. A wider distribution indicates that model bias is more variable over space and time. SC_{ij} distributions for each source were

stratified by season (Figs. D.1-D.4). For secondary organic carbon, sources with highly variable SC_{ij} values include biogenics (μ = 0.87 μg·m^{-3}, σ = 1.30 μg·m^{-3}) and solvents (μ = 0.19 μg·m^{-3}, σ = 0.32 μg·m^{-3}) (Fig. D.1). For nitrate, SC_{ij} values agricultural activities (μ = 0.084 μg·m^{-3}, σ = 0.81 μg·m^{-3}), natural gas combustion (e.g., industrial boilers, flares, heaters, and incinerators) (μ = 0.031 μg·m^{-3}, σ = 0.25 μg·m^{-3}), on-road diesel (μ = 0.012 μg·m^{-3}, σ = 0.24 μg·m^{-3}) and on-road gasoline (μ = 0.079 μg·m^{-3}, σ = 0.52 μg·m^{-3}) have high variability (Fig. D.2). There was also variability, to a lesser extent, in distributions for non-road diesel and gasoline. For ammonium SC_{ij} values, sources with high variability in distributions include agricultural activities (μ = 0.038 μg·m^{-3}, σ = 0.31 μg·m^{-3}) and coal combustion (μ = 0.025 μg·m^{-3}, σ = 0.26 μg·m^{-3}) (Fig. D.3). For sulfate SC_{ij} values, sources with highly varying distributions include coal combustion (μ = 0.27 μg·m^{-3}, σ = 1.28 μg·m^{-3}) and fuel oil combustion (μ = 0.045 μg·m^{-3}, σ = 0.23 μg·m^{-3}) (Fig. D.4).

5.3.3 Cross Validation

Withheld and spatially interpolated SC_{ij} values were compared at withheld site-day locations for each source and each secondary species (Table D5). For organic carbon, SC_{ij} values for agriculture/livestock had the highest correlation (r = 0.79) while fire had the lowest correlations (r = 0.02). Regression analysis of withheld and spatially interpolated SC_{ij} values led to regression intercepts near 0.00 with the exception of biogenics (α = 0.17) and solvents (α = 0.06). Regression slopes for each source ranged from 0.04 (fire) to 0.86 (agriculture/livestock). Normalized mean bias (NMB) was calculated for withheld and kriged SC_{ij} values:

$$NMB = \frac{\overline{SC_{ij}^{kriged}}}{\overline{SC_{ij}^{withheld}}} - 1 \qquad (5.7)$$

NMB was near zero for most sources, ranging from -0.05 (coal combustion) to 0.15 (fire) (Table D3).

For nitrate, the correlation coefficient was highest for agriculture/livestock (r = 0.64) and lowest for fire (r = 0.20). Regression intercepts were at or near zero for all sources. Regression slopes ranged from 0.41 (biogenics) to 1.20 (fuel oil combustion). NMB ranged from -10.9 (coal combustion) to 0.31 (dust), and the direction of the bias varied across the sources. For ammonium, the correlation coefficient was highest for coal combustion (r = 0.69) and lowest for fire (r = 0.07). The regression intercept was zero or near zero for all sources, and the regression slopes ranged from 0.16 (fire) to 0.88 (coal combustion). NMB ranged from -1.90 (non-road gasoline) to 0.14 (dust) and was mainly positive for most sources. For sulfate, the correlation coefficient was highest for coal combustion and on-road gasoline (r = 0.78) and lowest for biogenics (r = 0.04). The regression intercepts were near zero for all sources, and the regression slopes ranged from 0.06 (biogenics) to 1.01 (dust). NMB ranged from -0.45 (agriculture/livestock) to 2.32 (biogenics), and the direction of bias varied across the sources. Prediction performance was best for source/species pairs that contribute the majority of precursor species for secondary formation, such as biogenics and organic carbon, agriculture/livestock and nitrate/ammonium, and coal combustion and sulfate. Poorest performance was seen for sources with low contributions to a particular secondary species such as biogenic impacts on sulfate. Adequately capturing the temporal and spatial variability of some sources (e.g., dust and fires) in models is challenging and highly uncertain, which may also drive down prediction performance of the cross-validation application.

Concentrations of SANOC were reconstructed (Eq. 5.5) using withheld and kriged SC_{ij} values and were compared to observations to further evaluate the cross-validation (Table D6). Correlation coefficients for concentrations based on withheld SC_{ij} values reflected a near match to observations for all species, with correlations and regression slopes of 1.00 and NMB and regression intercepts of 0.00, though this is

expected given how the method is applied. Correlation coefficients for concentrations based on kriged SC_{ij} values were highest for sulfate (r_{krig} = 0.91) and lowest for organic carbon (r_{krig} = 0.69); hence, kriging did not capture the variability of withheld SC_{ij} values for organic carbon as well as for nitrate. Correlations did not reflect a near match when comparing concentrations calculated with withheld SC_{ij} values versus concentrations calculated with corresponding kriged SC_{ij} values. This is expected because 10% of the available data were removed for cross-validation. Regression intercepts were lowest for sulfate and ammonium (α_{krig} = 0.22) and highest for organic carbon and nitrate (α_{krig} = 0.43). Regression slopes were closest to one for sulfate (β_{krig} = 0.89) and furthest from one for nitrate (β_{krig} = 0.72).

Regression results indicate that the cross-validation performance of the secondary-adjustment is strongest for sulfate and weakest for organic carbon and nitrate. NMB was near zero for all species. Evaluation metrics for kriged comparisons did not degrade to a large extent, indicating fair agreement between predictions based on withheld and kriged SC_{ij} values. Organic carbon had consistently poorer performance metrics, potentially due to the TOT to TOR conversions that were applied to observed carbon data, which led to negative observations at times. This comparison highlights the better prediction performance of sulfate, however overall prediction performance for all species is favorable.

5.3.4 CMAQ-DDM Comparison

Performance of the secondary bias correction was further evaluated by comparing original CMAQ-DDM, hybrid, and secondary adjusted concentrations of SANOC and total $PM_{2.5}$ to observations (Table D.4, Fig. D.5). Mean concentrations for organic carbon over CSN sites-days were 2.22 ± 1.76, 1.61 ± 1.27, 1.08 ± 0.80, and 2.22 ± 1.76 $\mu g \cdot m^{-3}$ for observed, CMAQ-DDM, hybrid, and secondary-adjusted concentrations, respectively. The secondary-adjusted mean concentrations are the same as the observed mean, which is

much improved from CMAQ and hybrid concentrations. Mean concentrations for nitrate were 1.58 ± 2.29, 1.76 ± 2.55, 1.37 ± 2.12, and 1.58 ± 2.29 $\mu g \cdot m^{-3}$ for observed, CMAQ-DDM, hybrid, and newly adjusted concentrations, respectively. Again, the mean secondary-adjusted concentration is the same as the observation for nitrate. Mean concentrations for ammonium were 1.40 ± 1.18, 1.47 ± 1.22, 1.27 ± 1.04, and 1.40 ± 1.18 $\mu g \cdot m^{-3}$ for observed, CMAQ-DDM, hybrid, and newly adjusted concentrations, respectively. For sulfate, mean concentrations were 3.13 ± 2.61, 2.89 ± 2.30, 2.65 ± 2.12, and 3.13 ± 2.61 $\mu g \cdot m^{-3}$ for observed, CMAQ-DDM, hybrid, and secondary-adjusted concentrations, respectively. Mean concentrations of $PM_{2.5}$ were 12.6 ± 7.63, 12.6 ± 8.57, 8.52 ± 5.53, and 10.5 ± 6.26 $\mu g \cdot m^{-3}$ for the observation, CMAQ-DDM, hybrid, and secondary-adjustment, respectively. Mean CMAQ-DDM $PM_{2.5}$ concentration is very close to the mean observation, but CMAQ-DDM estimates of components remain biased without adjustment. Although secondary-adjusted SANOC concentrations match observations exactly for CSN locations, other modeled $PM_{2.5}$ components are still biased (e.g., non-carbon organic matter, metals, and unidentified $PM_{2.5}$ mass), leading the secondary-adjusted $PM_{2.5}$ concentration to not exactly match the observation.

The results show that the original CMAQ-DDM concentrations perform better for secondary species than the hybrid model from the preceding study (Ivey *et al*., 2015) [21]. It is important to note that the hybrid source apportionment methods adjusts concentrations and source contributions based upon primary species such as crustal and other fine metals, and did not address uncertainties in formation of SANOC species, which are the majority of the mass. This secondary bias correction was developed as an additional adjustment for the hybrid method to address this issue and the variation in secondary formation that is highly variable between locations, times, and species. More accurate PM species fields are developed by instituting both a primary and secondary adjustment to correct for modeled bias in concentrations and source impacts.

5.3.5 Source contributions (to secondary species)

The secondary bias correction was applied to improve source impacts on secondary species concentrations. The spatial extension also allows improved characterization of spatial and temporal variability of source impacts on secondary species concentrations. Spatial fields of seasonally-averaged source impacts on organic carbon are presented for biogenics, non-road gasoline, on-road gasoline, and solvents (Fig. D.6). Biogenic impacts on organic carbon are strongest in the southeastern and northwestern U.S. where vegetation is prevalent and denser than other regions. Impacts peak in the summer in the southeast (~4 $\mu g \cdot m^{-3}$ at peak). Non-road gasoline impacts on organic carbon are consistent throughout the year, and are highest in the eastern U.S. (~0.2 $\mu g \cdot m^{-3}$ at peak). On-road gasoline impacts on organic carbon are consistent throughout the year and are highest over urban centers across the U.S. (~0.4 $\mu g \cdot m^{-3}$ at peak). Solvent impacts on organic carbon are strongest in the winter and fall seasons and are greatest over the northeastern and western U.S. (~1 $\mu g \cdot m^{-3}$ at peak).

Spatial fields of seasonally-averaged source impacts on nitrate are presented for agriculture/livestock, natural gas combustion, on-road diesel, and on-road gasoline (Fig. D.7). Agriculture/livestock impacts on nitrate peak in the winter and fall seasons and are more significant in central California and over the Ohio River Valley in the Midwestern U.S. (~4 $\mu g \cdot m^{-3}$ at peak). Natural gas combustion impacts are constant throughout the year and are significant over central U.S. and southern California (~0.5 $\mu g \cdot m^{-3}$ at peak). On-road diesel impacts on nitrate are similar throughout the year and are greatest over the Ohio River Valley and southern California (~1 $\mu g \cdot m^{-3}$ at peak). On-road gasoline impacts on nitrate are also similar throughout the year (slightly higher in the winter) and are greatest over the eastern U.S. and southern California (~2 $\mu g \cdot m^{-3}$ at peak).

Spatial fields of seasonally-averaged source impacts on ammonium are presented for agriculture/livestock, coal combustion, fuel oil combustion, and on-road gasoline (Fig. D.8). Agriculture/livestock impacts on ammonium are constant throughout the year

and are most significant over the eastern U.S. and California (~1 $\mu g \cdot m^{-3}$ at peak). Coal combustion impacts on ammonium peak in the summer months and are most significant over the eastern U.S. (~1 $\mu g \cdot m^{-3}$ at peak). Fuel oil combustion impacts on ammonium are consistent throughout the year and are highest over southern California and southeastern U.S. (~0.5 $\mu g \cdot m^{-3}$ at peak). On-road gasoline impacts on ammonium are greatest in southern California and urban centers in the eastern U.S. and are consistent throughout the year (~0.5 $\mu g \cdot m^{-3}$ at peak).

Spatial fields of seasonally-averaged source impacts on sulfate are presented for coal combustion, fuel oil combustion, non-road others (non-road vehicles fueled by non-diesel/non-gasoline fuels), and other $PM_{2.5}$ sources (e.g., mineral processing, lawn waste burning, off-shore oil and gas activities, and marine vessels) (Fig. D.9). Coal combustion impacts on sulfate peak in the summer and are most significant over the eastern U.S. (~3 $\mu g \cdot m^{-3}$ at peak). Fuel oil combustion impacts on sulfate peak in the summer and are most significant in port areas such as New England, southern California, and the Gulf Coast (~1 $\mu g \cdot m^{-3}$ at peak). Non-road others' impacts on sulfate peak in the spring and summer months and are most significant along the coastal areas and in the Gulf of Mexico (~1 $\mu g \cdot m^{-3}$ at peak). Other $PM_{2.5}$ sources' impacts on sulfate are most significant in the southern U.S. and northern mountain region and peak in the spring and summer months (~1 $\mu g \cdot m^{-3}$ at peak).

Seasonal source contributions for SANOC and total $PM_{2.5}$ are highlighted for three U.S. cities, Atlanta, Georgia; Los Angeles, California; and Pittsburgh, Pennsylvania (see map in Fig. 5.1). Atlanta is an inland city located in southeastern U.S., a heavily forested region with high emissions from biogenic sources. The conjunction of three major U.S. interstates occurs in Atlanta, and the busiest corridor sees up to 300,000 vehicles per day [102]. Electricity needs in the metro area are served by coal- and natural gas-fired power plants. The results from this study reflect the characteristics of Atlanta emissions activity. The Atlanta airport is also the busiest operations airport in the world

with almost 2,500 arrivals and departures daily, which contributes to emissions of secondary precursors. In Atlanta, organic carbon contributions are mainly from biogenics, on-road sources, and solvents, and biogenics peak in the summer (2.51μg·m^{-3}) (Table 5.1). Nitrate is mostly contributed by agriculture/livestock and on-road sources and both contributions peak in the colder seasons (1 μg·m^{-3}). Ammonium is contributed by agriculture/livestock, coal combustion, and on-road sources and contributions for the sources peak in the warmer months. Impacts from sulfate originate from coal (summer peak of 4.61 μg·m^{-3}) and fuel oil and diesel combustion sources, and all sources peak in the summer months. In Atlanta, agriculture/livestock, biogenics, coal combustion, and mobile sources dominate total $PM_{2.5}$ contributions (Table D7).

Los Angeles is located near the coast of southern California, and its local climate is influenced by its proximity to the Pacific Ocean as well as its positioning between the ocean to the west and several small mountain ranges to the north and east. The topography of the Los Angeles area leads to temperature inversions and air getting trapped by the mountains. Los Angeles is a city with large urban sprawl, leading to high daily traffic volumes. With several ports along the California coast, shipping emissions from fuel oil burning contribute to coastal aerosol concentrations. Presently, the majority of California's energy is produced by natural gas and renewable fuels [103103]. In previous years, hydroelectric and nuclear power supplied close to one-third of California's energy. The nearby San Joaquin Valley is a highly polluted area of California due to emissions from dairy and produce farms. In Los Angeles, organic carbon was contributed by biogenics (winter peak of 2.41 μg·m^{-3}), meat cooking, on-road gasoline, and solvents (Table 5.2). Nitrate was mainly contributed by on-road gasoline (winter peak of 2.36 μg·m^{-3}), agriculture/livestock (winter peak of 1.09 μg·m^{-3}), and natural gas combustion. Ammonium was mostly contributed by agriculture/livestock, fuel oil and natural gas combustion, and on-road gasoline vehicles (spring peak of 0.67 μg·m^{-3}). Sulfate was mostly contributed by fuel oil combustion (spring peak of 0.56

$\mu g \cdot m^{-3}$) and non-road others. In Los Angeles, agriculture/livestock, biogenics, mobile sources, and fuel oil and natural gas combustion dominate total $PM_{2.5}$ contributions (Table D8).

Pittsburgh is an inland city, once characterized by its heavily industrial economy, where it was home to more than 300 steel-related businesses and coal production facilities. Today, the economy is service- and higher education-based, and most electricity is generated with natural gas and nuclear power, while coal production still persists [104]. In Pittsburgh, organic carbon was mainly contributed by biogenics (summer peak of 1.55 $\mu g \cdot m^{-3}$), on-road gasoline vehicles, and solvents (Table 5.3). Nitrate originated from agriculture/livestock (winter peak of 1.62 $\mu g \cdot m^{-3}$), on-road gasoline vehicles (winter peak of 0.67 $\mu g \cdot m^{-3}$), and other $PM_{2.5}$ sources. Ammonium originated from agriculture/livestock (summer peak of 0.83 $\mu g \cdot m^{-3}$), coal combustion (summer peak of 0.76 $\mu g \cdot m^{-3}$), and on-road gasoline vehicles. Sulfate was mainly contributed by coal (winter peak of 6.00 $\mu g \cdot m^{-3}$) and agriculture/livesrock (summer peak of 0.83 $\mu g \cdot m^{-3}$). In Pittsburgh, agriculture/livestock, coal gas combustion, and mobile sources dominate total $PM_{2.5}$ contributions (Table D9).

Table 5.1 Source impacts on secondary $PM_{2.5}$ for Atlanta, GA. Top 5 impacting sources in 2006 based on new adjustments for organic carbon (OC), nitrate (NO_3), ammonium (NH_4), and sulfate (SO_4) for Atlanta, GA and source impacts in $\mu g \cdot m^{-3}$. Source impacts and total observed $PM_{2.5}$ are seasonally averaged. Sources (abbreviations) include: agricultural activities and livestock operations (ag), aircraft (air), biogenics (biog), coal combustion (coal), dust, fires, fuel oil combustion (foil), meat cooking (meat), metals processing (metal), natural gas combustion (ng), non-road diesel (nrd), non-road gasoline (nrg), non-road others (nro), on-road diesel (ord), on-road gasoline (org), others (ot), other combustion (otc), solvents (slv), sea salt (ss), wood burning (wood).

species	**rank**	**winter ($\mu g \cdot m^{-3}$)**		**spring ($\mu g \cdot m^{-3}$)**		**summer ($\mu g \cdot m^{-3}$)**		**fall ($\mu g \cdot m^{-3}$)**	
$PM_{2.5}$	-	11.8		15.0		22.3		15.3	
OC	1	1.17	biog	1.54	biog	2.51	biog	1.83	biog
	2	0.30	slv	0.22	ot	0.24	ord	0.25	slv
	3	0.29	org	0.19	org	0.24	ot	0.22	org
	4	0.22	ord	0.18	ord	0.20	org	0.20	ord
	5	0.19	ot	0.15	slv	0.15	nrg	0.17	ot
NO3	1	1.09	ag	0.78	ag	0.38	org	0.94	ag
	2	0.96	org	0.41	org	0.34	ag	0.63	org
	3	0.16	ord	0.15	ord	0.14	ord	0.17	ord
	4	0.10	ot	0.09	nrd	0.07	nrd	0.10	nrd
	5	0.08	nrd	0.05	ot	0.06	coal	0.06	ot
NH4	1	0.37	ag	0.52	coal	0.98	coal	0.57	coal
	2	0.28	org	0.47	ag	0.44	ag	0.44	ag
	3	0.19	coal	0.23	org	0.40	org	0.30	org
	4	0.04	ot	0.09	ord	0.15	ord	0.09	ord
	5	0.04	ord	0.07	foil	0.11	foil	0.07	foil
SO4	1	0.83	coal	2.44	coal	4.61	coal	2.30	coal
	2	0.13	foil	0.28	foil	0.44	foil	0.24	foil
	3	0.04	ot	0.13	ord	0.42	ord	0.13	ord
	4	0.01	nro	0.12	ot	0.25	org	0.07	ot
	5	0.01	metal	0.08	nrd	0.21	nrd	0.06	nrd

Table 5.2. Source impacts on secondary $PM_{2.5}$ for Los Angeles, CA. Top 5 impacting sources in 2006 based on new adjustments for organic carbon (OC), nitrate (NO_3), ammonium (NH_4), and sulfate (SO_4) for Los Angeles, CA and source impacts in $\mu g \cdot m^{-3}$. Source impacts and total observed $PM_{2.5}$ are seasonally averaged. Sources (abbreviations) include: agricultural activities and livestock operations (ag), aircraft (air), biogenics (biog), coal combustion (coal), dust, fires, fuel oil combustion (foil), meat cooking (meat), metals processing (metal), natural gas combustion (ng), non-road diesel (nrd), non-road gasoline (nrg), non-road others (nro), on-road diesel (ord), on-road gasoline (org), others (ot), other combustion (otc), solvents (slv), sea salt (ss), wood burning (wood).

species	**rank**	**winter ($\mu g \cdot m^{-3}$)**		**spring ($\mu g \cdot m^{-3}$)**		**summer ($\mu g \cdot m^{-3}$)**		**fall ($\mu g \cdot m^{-3}$)**	
$PM_{2.5}$		17.2		12.8		16.8		15.2	
OC	1	2.41	biog	0.81	biog	0.85	biog	1.68	biog
	2	0.61	slv	0.33	meat	0.29	meat	0.29	meat
	3	0.26	meat	0.31	slv	0.27	org	0.28	org
	4	0.22	org	0.21	org	0.19	nrd	0.27	slv
	5	0.14	wood	0.15	nrd	0.12	slv	0.19	nrd
NO_3	1	2.36	org	2.22	org	1.78	org	1.77	org
	2	1.09	ag	1.38	ag	0.96	ng	0.91	ng
	3	0.97	ng	1.29	ng	0.66	ag	0.91	ag
	4	0.74	ord	0.58	nrd	0.35	nrd	0.47	nrd
	5	0.72	foil	0.49	ord	0.32	nro	0.45	ord
NH_4	1	0.64	foil	0.67	org	0.46	org	0.44	org
	2	0.59	org	0.42	ng	0.36	nro	0.35	foil
	3	0.28	ng	0.40	ag	0.26	ng	0.25	ng
	4	0.22	ag	0.29	foil	0.24	foil	0.23	ag
	5	0.18	ord	0.22	nrd	0.20	nrd	0.15	nrd
SO_4	1	0.51	foil	0.56	foil	0.70	nro	0.46	foil
	2	0.10	ot	0.31	nro	0.51	foil	0.14	nro
	3	0.07	coal	0.20	ot	0.39	nrd	0.12	ot
	4	0.04	nro	0.17	nrd	0.31	ot	0.07	nrd
	5	0.02	ng	0.09	coal	0.17	coal	0.07	coal

Table 5.3. Source impacts on secondary $PM_{2.5}$ for Pittsburgh, PA. Top 5 impacting sources in 2006 based on new adjustments for organic carbon (OC), nitrate (NO_3), ammonium (NH_4), and sulfate (SO_4) for Pittsburgh, PA and source impacts in $\mu g \cdot m^{-3}$. Source impacts and total observed $PM_{2.5}$ are seasonally averaged. Sources (abbreviations) include: agricultural activities and livestock operations (ag), aircraft (air), biogenics (biog), coal combustion (coal), dust, fires, fuel oil combustion (foil), meat cooking (meat), metals processing (metal), natural gas combustion (ng), non-road diesel (nrd), non-road gasoline (nrg), non-road others (nro), on-road diesel (ord), on-road gasoline (org), others (ot), other combustion (otc), solvents (slv), sea salt (ss), wood burning (wood).

species	**rank**	**winter ($\mu g \cdot m^{-3}$)**		**spring ($\mu g \cdot m^{-3}$)**		**summer ($\mu g \cdot m^{-3}$)**		**fall ($\mu g \cdot m^{-3}$)**	
$PM_{2.5}$		11.0		12.0		12.0		10.3	
OC	1	0.37	biog	0.59	biog	1.55	biog	0.77	biog
	2	0.24	slv	0.18	slv	0.19	slv	0.30	slv
	3	0.22	org	0.18	org	0.18	meat	0.18	org
	4	0.12	meat	0.17	meat	0.17	nrd	0.16	meat
	5	0.10	wood	0.14	wood	0.17	nrg	0.12	nrd
NO_3	1	1.62	ag	1.20	ag	0.65	ag	1.21	ag
	2	0.67	org	0.47	org	0.35	org	0.51	org
	3	0.18	ot	0.15	ot	0.10	ord	0.14	ot
	4	0.12	ord	0.13	ord	0.08	ot	0.09	ord
	5	0.11	mg	0.07	nrd	0.08	nrd	0.07	nrd
NH_4	1	0.51	ag	0.57	ag	0.83	ag	0.57	ag
	2	0.23	org	0.32	coal	0.76	coal	0.30	coal
	3	0.21	coal	0.22	org	0.35	org	0.27	org
	4	0.17	ot	0.19	ot	0.26	ot	0.18	ot
	5	0.03	foil	0.04	ord	0.10	ord	0.03	foil
SO_4	1	1.29	coal	2.26	coal	6.00	coal	1.98	coal
	2	0.13	foil	0.13	ot	0.31	ord	0.11	foil
	3	0.06	ot	0.13	foil	0.28	ot	0.07	ot
	4	0.05	metal	0.06	metal	0.28	nrd	0.06	metal
	5	0.02	nrg	0.02	otc	0.19	foil	0.01	ord

5.4 Discussion

Overall uncertainty in secondary concentrations and source impacts primarily results from uncertainties in the emissions, chemistry, gas/particle partitioning, and meteorology. In this case, uncertainty in modeled concentrations can be investigated by examining the relationship between observed and predicted concentrations and SC_{ij} values. For example, the ratios of mean organic carbon SC_{ij} for biogenics, on-road gasoline vehicles, and solvents to observed organic carbon for CSN sites are 0.39, 0.01, and 0.09, respectively. In other words, modeled impacts on secondary organic carbon are underestimated in CMAQ by approximately 39% from biogenics, 1% from on-road gasoline vehicles, and 9% from solvents on average. The ratios of agriculture/livestock, natural gas combustion, on-road diesel, and on-road gasoline SC_{ij} to observed nitrate are 0.05, 0.02, 0.007, and 0.05, respectively. This indicates an approximate underestimation of impact on nitrate by 5% for agriculture/livestock, 2.0% for natural gas combustion, 0.7% for on-road diesel vehicles, and 5.0% for on-road gasoline vehicles. The ratios of agriculture/livestock, coal combustion, fuel oil combustion, and on-road gasoline SC_{ij} to observed ammonium are 0.027, 0.02, 0.014, and 0.015, respectively. This indicates an approximate underestimation of impact on ammonium by 2.7% for agriculture/livestock, 2.0% for coal combustion, 1.4% for fuel oil combustion, and 1.5% for on-road gasoline vehicles. The ratios of coal combustion, fuel oil combustion, and on-road diesel SC_{ij} to observed sulfate are 0.09, 0.014, and 0.007, respectively. This indicates an approximate underestimation of impact on sulfate by 9% for coal combustion, 1.4% for fuel oil combustion, and 0.7% for on-road diesel. Ratios can represent the approximate bias in modeled source impacts and concentrations. Although the magnitude of the bias is small for some source/species pairs, the cumulative effect of bias over all sources is significant.

Examining seasonal and spatial variability of SC_{ij} values lends information about the spatial and temporal uncertainty of modeled secondary $PM_{2.5}$. For secondary organic

carbon (SOC), biogenic impacts are underestimated near the northwest U.S. border in cooler months, which is portrayed by high SC_{ij} values in the region (Fig. 5.2). This uncertainty may be due to the choice of emission factors from vegetation or wintertime partitioning of volatile organic carbon in the region. Gasoline impacts on SOC are underestimated in the southeast and near the northwest U.S. border throughout the year. For nitrate, impacts from most sources are underestimated in central/southern California throughout the year (Fig. 5.3). Low nitrate impacts may stem from uncertainties in chemical titrations in the region or underestimations of primary emissions of nitrogen oxides. For ammonium, impact from on-road gasoline is underestimated in the southeast in the summer and underestimated throughout the year in southern California (Fig 5.4). In cooler months, agriculture/livestock and fuel oil impacts on ammonium are underestimated in southern California. Low ammonium impacts may also stem from uncertainties in chemical titrations in the region as well as uncertainties in emissions estimates of ammonia. Impacts from coal combustion are underestimated in the summer in the southeast. For sulfate, coal combustion impacts are underestimated in the northeast in winter and underestimated over the eastern U.S. in spring and summer (Fig. 5.5). Uncertainty in modeling temperature-dependent phase partitioning sulfate may affect impacts from coal combustion, as bias in emissions is less likely due to continuous monitoring of stack emissions of sulfur dioxide [105]. Impacts from off-road and off-shore sources are underestimated in the southern coastal areas, and may be due to uncertainties in emissions from the sources.

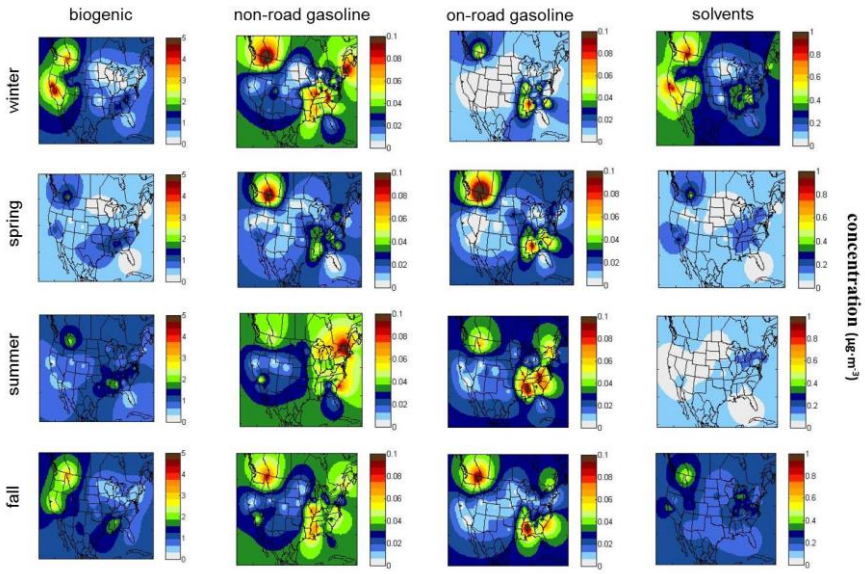

Figure 5.2: Seasonally-averaged spatial fields of $\boldsymbol{SC_{ij}}$ values in μg·m^{-3} for top secondary organic carbon sources.

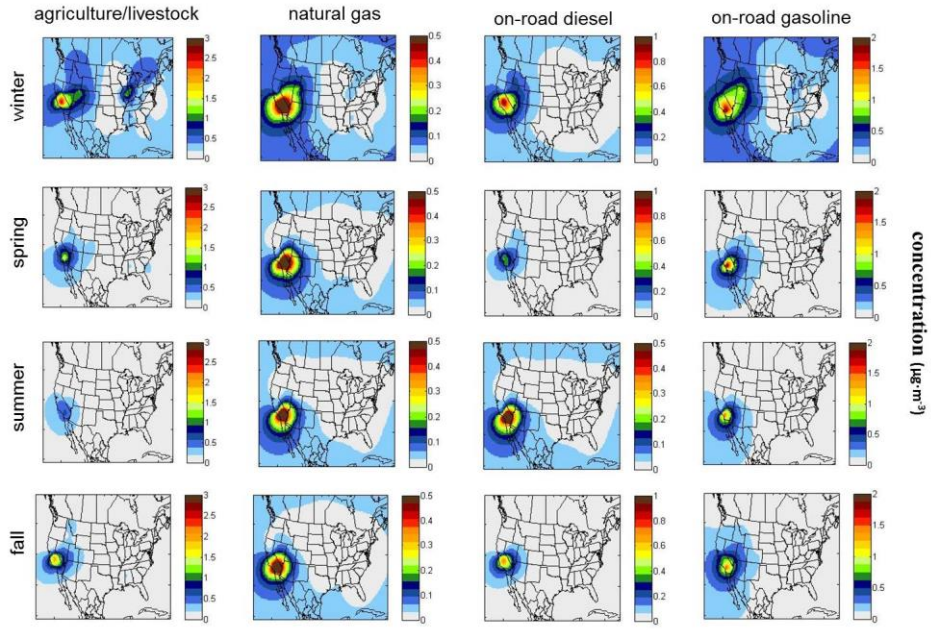

Figure 5.3. Seasonally-averaged spatial fields of $\boldsymbol{SC_{ij}}$ values in μg·m^{-3} for top nitrate sources.

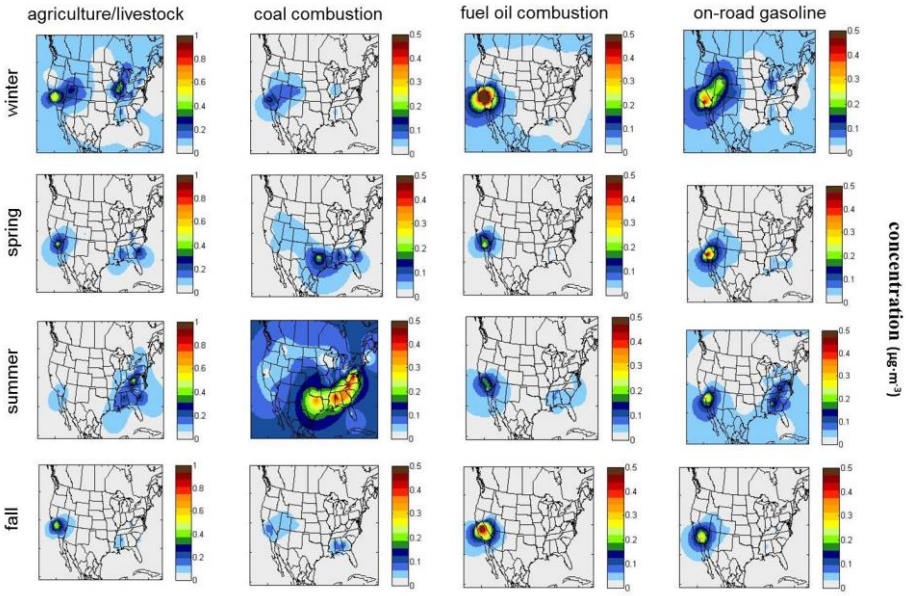

Figure 5.4. Seasonally-averaged spatial fields of $\boldsymbol{SC_{ij}}$ values in μg·m^{-3} for top ammonium sources.

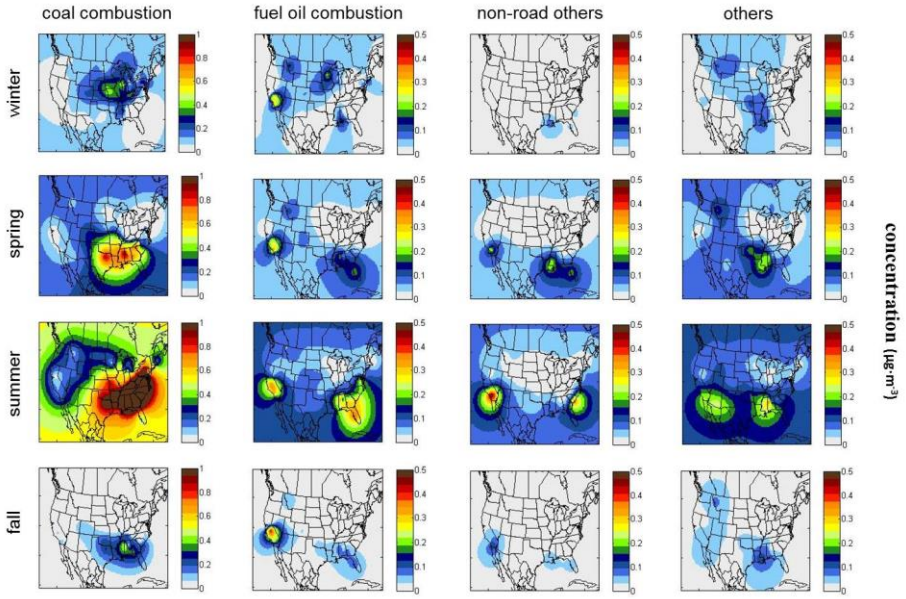

Figure 5.5: Seasonally-averaged spatial fields of $\boldsymbol{SC_{ij}}$ values in μg·m^{-3} for top sulfate sources.

5.5 Conclusion

Newly adjusted concentrations better agree with observed concentrations compared to original CMAQ-DDM (base case) estimates before adjustment. The adjustments are also applied to modeled source impacts, which improve estimates of impacts on secondary species. Results from this analysis imply that, for the modeling scheme used here, OC is underestimated by approximately 49%, nitrate by 7.7%, ammonium by 7.6%, and sulfate by 11.1%. The methods presented here are novel in that traditional receptor modeling analyses do not partition out the secondary source impacts. The secondary bias correction can be applied for other modeling schemes (e.g. other versions of CMAQ or different meteorological and emissions inputs), to determine overall bias in secondary $PM_{2.5}$ concentrations. This study also illuminates the greatest impacting sources on each secondary species over the continental U.S., which helps to identify the modeling pathways with most uncertainty.

5.6 Acknowledgements

This publication was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901, STAR Fellowship FP-91761401-0, and by NASA under grant NNX11AI55G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Alfred P. Sloan Foundation for their support.

CHAPTER 6: CONCLUSION

The methods presented in this dissertation improve our understanding of $PM_{2.5}$ source impacts over continental U.S. (CONUS), identify potential sources of bias in current methods, and increase source impact modeling strengths and capabilities. The methods are novel data assimilation techniques that combine both observed and simulated data to generate better estimates of $PM_{2.5}$ sources impacts and concentrations. The results presented offer insights into the greatest impacting sources of $PM_{2.5}$ at a national scale over longer time periods, which is the first demonstration of this effort as revealed from the current literature. The methods and results are beneficial for policy and health applications, such as establishing NAAQAS attainment for municipalities and further exploring links between human exposure to $PM_{2.5}$ sources and the development or worsening of adverse health effects. The major conclusions drawn from the previous chapters are briefly reiterated below.

The spatial hybrid (SH) model is an effective approach for improving simulated source impact spatial fields through statistical optimization instead of re-running CMAQ–DDM, which is more computationally expensive. The method combines 1-in-3 day observations at monitoring locations and daily, 36-km resolution CMAQ-DDM source impacts, along with measurement and model uncertainties, to generate optimized spatial fields of source impacts that better reflect observations. The application and evaluation of the spatial hybrid approach for January 2004 showed that the method was successful in improving simulated concentrations of trace metal species such as silicon, potassium, iron, and aluminum. Reducing biases in modeled concentrations of trace metals improved estimates of source impacts, as tracer species are often critical for distinguishing the impacts from sources with similar profiles in traditional source apportionment approaches (e.g. the chemical mass balance method). The SH approach also allows the impacts from sources that primarily emit precursors of secondary $PM_{2.5}$

species to be attributed directly to the source, specifically coal combustion (sulfur dioxide), agricultural/livestock activities (ammonia), and biogenics (volatile organic compounds), instead of being attributed to gas-to-particle interactions. This SH approach can be applied over long time periods to produce a historical record of source impacts in the United States for use in long-term health analyses.

The SH method was applied for three years, 2005 to 2007, to assess seasonal and regional variability in source impacts for 16 source categories over the CONUS. Each year of CMAQ-DDM simulations were generated using different modeling schemes, leading to different emissions patterns for some sources. Distributions of hybrid-optimized adjustment factors varied for each source year-to-year. The greatest impacting sources identified in this multi-year assessment were coal combustion, on-road vehicles, and agriculture/livestock activities for 2005 and 2006, and fire, dust, and on-road vehicles in 2007. An urban-rural assessment was conducted to determine spatial gradients in three U.S. cities. Results indicated that while source contributions are similar in both urban and rural locations, mobile vehicle impacts were a higher percentage in urban areas while dust impacts were a higher in the rural locations. Additional assessment was conducted by comparing results for three U.S. regions for which nonattainment studies had been conducted. The SH method results were comparable to the peer-reviewed studies, demonstrating that SH methods can be useful for identifying $PM_{2.5}$ sources that impact a region's nonattainment status. Furthermore, the SH approach quantifies source impacts with higher source stratification compared with receptor modeling approaches. Because the method is source-oriented, the SH method can potentially identify previously unrecognized sources that impact a region through long-range transport, such as dust and fires, by examining upwind emissions along with local source impacts.

A nonlinear optimization method for revising reference $PM_{2.5}$ source profiles is developed and applied evaluated over all available CSN monitors for the year 2006. The method takes into account both observations and modeled source impact estimates to

generate receptor-based source profiles for 20 source categories and 23 $PM_{2.5}$ component species. Analysis of the revised profiles reveals that coal combustion, metals processing, and sea salt profiles underwent the most change over all CSN monitors, in comparison to the reference profile. Further analysis revealed that bias in modeled concentrations of trace metals highly influenced the profile optimization, where the new profiles correct for this bias and improve modeled source impact estimates. The revised profiles were further evaluated by implementing the profiles in the CMB-iteration model. Profiles for seven common sources were used to calculate source impacts for Atlanta, GA and St. Louis, MO (mobile sources (gasoline and diesel), dust, biomass burning, coal combustion or metals processing, sulfate, nitrate, and ammonium) models. The CMB analysis revealed that the proper selection of profiles is critical for producing reasonable source impacts. For example the inclusion of the metals processing profile in the Atlanta analysis led to negative contributions from the metals source, which led to its exclusion from the analysis. Additionally, most CMB-generated source impacts were comparable to SH impacts. This data-integrated approach for $PM_{2.5}$ source profile optimization is beneficial for source impact analysis in regions where reference profiles have high uncertainty or are unknown. The incorporation of observations allows the revised profiles to reflect local pollution conditions.

A method for quantifying bias in modeled secondary $PM_{2.5}$ is developed and applied over CONUS for the year 2006. The method distributes the bias between observations and SH concentrations of sulfate, nitrate, ammonium, and secondary organic carbon (SANOC) across source impacts for 20 source categories. The bias adjustments provide information regarding which species/source relationships are subject to the greatest bias. The 2006 analysis determined that biogenics, solvents, and on-road gasoline impacts contributed to the majority of the bias in modeled secondary organic carbon concentrations; agriculture/livestock and on-road sources for nitrate; agriculture/livestock and on-road sources for ammonium; and coal and fuel oil

combustion for sulfate. Additionally, secondary-adjusted concentrations of SANOC and $PM_{2.5}$ better agree with observed concentrations compared to original CMAQ-DDM (base case) estimates before adjustment. This study illuminates the greatest impacting sources on each secondary species over continental U.S., which helps to identify the modeling pathways with most uncertainty.

6.1 Future Work

The work presented in this dissertation is a foundation for additional research in the area of $PM_{2.5}$ modeling and source apportionment. The SH method can be applied for multiple years (2005-2012), and the results can be made available in a public database. This database can provide a consistent set of data for researchers to use in several fields, such as air quality, public policy, city planning, biostatistics, and epidemiology. The modeling uncertainties highlighted in the secondary analysis can serve as a guide to improving modeling mechanisms secondary $PM_{2.5}$. The novel methods presented here can also be applied to gaseous pollutants such as nitrogen oxides and ozone to improve estimates of source contributions to ambient concentrations. In Figure 6.1, monthly-averaged, first-order sensitivities of ozone to emissions of on-road gasoline and diesel vehicles and coal combustion are presented for May – August 2007 (peak of the ozone season). With the tightening of the NAAQS for ozone, determining the greatest impacting sources of ozone precursors will be essential for attainment efforts for affected municipalities. In summary, data assimilation methods are beneficial for air quality modeling assessments because the combined strengths of observations and models reduce the inherent biases and uncertainties of using solely one source of data.

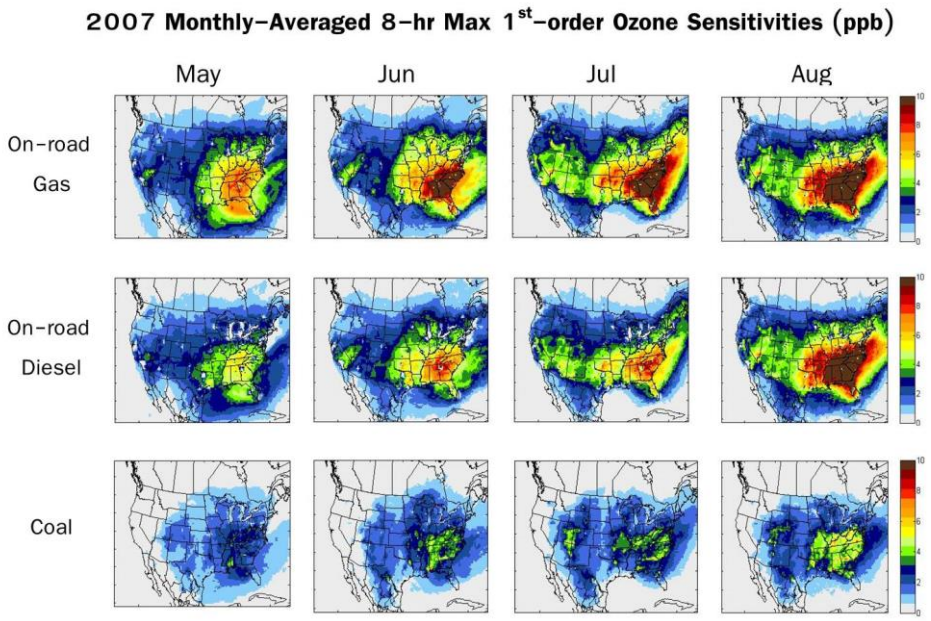

Figure 6.1. Monthly-averaged summertime 8-hr maximum first order ozone sensitivities (ppb) for continental U.S.

REFERENCES

1. Guttikunda, S. K.; Calori, G., A GIS based emissions inventory at 1 km x 1 km spatial resolution for air pollution analysis in Delhi, India. *Atmospheric Environment* 2013, *67*, 101-111.

2. Ji, D. S.; Li, L.; Wang, Y. S.; Zhang, J. K.; Cheng, M. T.; Sun, Y.; Liu, Z. R.; Wang, L. L.; Tang, G. Q.; Hu, B.; Chao, N.; Wen, T. X.; Miao, H. Y., The heaviest particulate air-pollution episodes occurred in northern China in January, 2013: Insights gained from observation. *Atmospheric Environment* 2014, *92*, 546-556.

3. Fann, N.; Lamson, A. D.; Anenberg, S. C.; Wesson, K.; Risley, D.; Hubbell, B. J., Estimating the National Public Health Burden Associated with Exposure to Ambient PM2.5 and Ozone. *Risk Anal* 2012, *32*, (1), 81-95.

4. Kunzli, N.; Jerrett, M.; Mack, W. J.; Beckerman, B.; LaBree, L.; Gilliland, F.; Thomas, D.; Peters, J.; Hodis, H. N., Ambient air pollution and atherosclerosis in Los Angeles. *Environ Health Persp* 2005, *113*, (2), 201-206.

5. Kappos, A. D.; Bruckmann, P.; Eikmann, T.; Englert, N.; Heinrich, U.; Hoppe, P.; Koch, E.; Krause, G. H. M.; Kreyling, W. G.; Rauchfuss, K.; Rombout, P.; Schulz-Klemp, V.; Thiel, W. R.; Wichmann, H. E., Health effects of particles in ambient air. *Int J Hyg Envir Heal* 2004, *207*, (4), 399-407.

6. Lim, S. S.; Vos, T.; Flaxman, A. D.; Danaei, G.; Shibuya, K.; Adair-Rohani, H.; Amann, M.; Anderson, H. R.; Andrews, K. G.; Aryee, M.; Atkinson, C.; Bacchus, L. J.; Bahalim, A. N.; Balakrishnan, K.; Balmes, J.; Barker-Collo, S.; Baxter, A.; Bell, M. L.; Blore, J. D.; Blyth, F.; Bonner, C.; Borges, G.; Bourne, R.; Boussinesq, M.; Brauer, M.; Brooks, P.; Bruce, N. G.; Brunekreef, B.; Bryan-Hancock, C.; Bucello, C.; Buchbinder, R.; Bull, F.; Burnett, R. T.; Byers, T. E.; Calabria, B.; Carapetis, J.; Carnahan, E.; Chafe, Z.; Charlson, F.; Chen, H. L.; Chen, J. S.; Cheng, A. T. A.; Child, J. C.; Cohen, A.; Colson, K. E.; Cowie, B. C.; Darby, S.; Darling, S.; Davis, A.; Degenhardt, L.; Dentener, F.; Des Jarlais, D. C.; Devries, K.; Dherani, M.; Ding, E. L.; Dorsey, E. R.; Driscoll, T.; Edmond, K.; Ali, S. E.; Engell, R. E.; Erwin, P. J.; Fahimi, S.; Falder, G.; Farzadfar, F.; Ferrari, A.; Finucane, M. M.; Flaxman, S.; Fowkes, F. G. R.; Freedman, G.; Freeman, M. K.; Gakidou, E.; Ghosh, S.; Giovannucci, E.; Gmel, G.; Graham, K.; Grainger, R.; Grant, B.; Gunnell, D.; Gutierrez, H. R.; Hall, W.; Hoek, H. W.; Hogan, A.; Hosgood, H. D.; Hoy, D.; Hu, H.; Hubbell, B. J.; Hutchings, S. J.; Ibeanusi, S. E.; Jacklyn, G. L.; Jasrasaria, R.; Jonas, J. B.; Kan, H. D.; Kanis, J. A.; Kassebaum, N.; Kawakami, N.; Khang, Y. H.; Khatibzadeh, S.; Khoo, J. P.; Kok, C.; Laden, F.; Lalloo, R.; Lan, Q.; Lathlean, T.; Leasher, J. L.; Leigh, J.; Li, Y.; Lin, J. K.; Lipshultz, S. E.; London, S.; Lozano, R.; Lu, Y.; Mak, J.; Malekzadeh, R.; Mallinger, L.; Marcenes, W.; March, L.; Marks, R.; Martin, R.; McGale, P.; McGrath, J.; Mehta, S.; Mensah, G. A.; Merriman, T. R.; Micha, R.; Michaud, C.;

Mishra, V.; Hanafiah, K. M.; Mokdad, A. A.; Morawska, L.; Mozaffarian, D.; Murphy, T.; Naghavi, M.; Neal, B.; Nelson, P. K.; Nolla, J. M.; Norman, R.; Olives, C.; Omer, S. B.; Orchard, J.; Osborne, R.; Ostro, B.; Page, A.; Pandey, K. D.; Parry, C. D. H.; Passmore, E.; Patra, J.; Pearce, N.; Pelizzari, P. M.; Petzold, M.; Phillips, M. R.; Pope, D.; Pope, C. A.; Powles, J.; Rao, M.; Razavi, H.; Rehfuess, E. A.; Rehm, J. T.; Ritz, B.; Rivara, F. P.; Roberts, T.; Robinson, C.; Rodriguez-Portales, J. A.; Romieu, I.; Room, R.; Rosenfeld, L. C.; Roy, A.; Rushton, L.; Salomon, J. A.; Sampson, U.; Sanchez-Riera, L.; Sanman, E.; Sapkota, A.; Seedat, S.; Shi, P. L.; Shield, K.; Shivakoti, R.; Singh, G. M.; Sleet, D. A.; Smith, E.; Smith, K. R.; Stapelberg, N. J. C.; Steenland, K.; Stockl, H.; Stovner, L. J.; Straif, K.; Straney, L.; Thurston, G. D.; Tran, J. H.; Van Dingenen, R.; van Donkelaar, A.; Veerman, J. L.; Vijayakumar, L.; Weintraub, R.; Weissman, M. M.; White, R. A.; Whiteford, H.; Wiersma, S. T.; Wilkinson, J. D.; Williams, H. C.; Williams, W.; Wilson, N.; Woolf, A. D.; Yip, P.; Zielinski, J. M.; Lopez, A. D.; Murray, C. J. L.; Ezzati, M., A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012, *380*, (9859), 2224-2260.

7. Dockery, D. W.; Pope, C. A.; Xu, X. P.; Spengler, J. D.; Ware, J. H.; Fay, M. E.; Ferris, B. G.; Speizer, F. E., An Association between Air-Pollution and Mortality in 6 United-States Cities. *New England Journal of Medicine* 1993, *329*, (24), 1753-1759.

8. Yue, H. F.; Yun, Y.; Gao, R.; Li, G. K.; Sang, N., Winter Polycyclic Aromatic Hydrocarbon-Bound Particulate Matter from Peri-urban North China Promotes Lung Cancer Cell Metastasis. *Environ Sci Technol* 2015, *49*, (24), 14484-14493.

9. Du, Y. X.; Xu, X. H.; Chu, M.; Guo, Y.; Wang, J. H., Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. *J Thorac Dis* 2016, *8*, (1), E8-E19.

10. Wang, F.; Ni, S. S.; Liu, H., Pollutional haze and COPD: etiology, epidemiology, pathogenesis, pathology, biological markers and therapy. *J Thorac Dis* 2016, *8*, (1), E20-E30.

11. Liu, S. K.; Cai, S.; Chen, Y.; Xiao, B.; Chen, P.; Xiang, X. D., The effect of pollutional haze on pulmonary function. *J Thorac Dis* 2016, *8*, (1), E41-E56.

12. Akintoye, E.; Shi, L. H.; Obaitan, I.; Olusunmade, M.; Wang, Y.; Newman, J. D.; Dodson, J. A., Association between fine particulate matter exposure and subclinical atherosclerosis: A meta-analysis. *Eur J Prev Cardiol* 2016, *23*, (6), 602-612.

13. Bates, J. T.; Weber, R. J.; Abrams, J.; Verma, V.; Fang, T.; Klein, M.; Strickland, M. J.; Sarnat, S. E.; Chang, H. H.; Mulholland, J. A.; Tolbert, P. E.; Russell, A. G., Reactive Oxygen Species Generation Linked to Sources of Atmospheric

Particulate Matter and Cardiorespiratory Effects. *Environ Sci Technol* 2015, *49*, (22), 13605-13612.

14. Gass, K.; Balachandran, S.; Chang, H. H.; Russell, A. G.; Strickland, M. J., Ensemble-Based Source Apportionment of Fine Particulate Matter and Emergency Department Visits for Pediatric Asthma. *Am J Epidemiol* 2015, *181*, (7), 504-512.

15. Gass, K.; Klein, M.; Sarnat, S. E.; Winquist, A.; Darrow, L. A.; Flanders, W. D.; Chang, H. H.; Mulholland, J. A.; Tolbert, P. E.; Strickland, M. J., Associations between ambient air pollutant mixtures and pediatric asthma emergency department visits in three cities: a classification and regression tree approach. *Environ Health-Glob* 2015, *14*.

16. Pearce, J. L.; Waller, L. A.; Mulholland, J. A.; Sarnat, S. E.; Strickland, M. J.; Chang, H. H.; Tolbert, P. E., Exploring associations between multipollutant day types and asthma morbidity: epidemiologic applications of self-organizing map ambient air quality classifications. *Environ Health-Glob* 2015, *14*.

17. Friberg, M. D.; Zhai, X. X.; Holmes, H. A.; Chang, H. H.; Strickland, M. J.; Sarnat, S. E.; Tolbert, P. E.; Russell, A. G.; Mulholland, J. A., Method for Fusing Observational Data and Chemical Transport Model Simulations To Estimate Spatiotemporally Resolved Ambient Air Pollution. *Environ Sci Technol* 2016, *50*, (7), 3695-3705.

18. Sturtz, T. M.; Schichtel, B. A.; Larson, T. V., Coupling Chemical Transport Model Source Attributions with Positive Matrix Factorization: Application to Two IMPROVE Sites Impacted by Wildfires. *Environ Sci Technol* 2014, *48*, (19), 11389-11396.

19. Viana, M.; Kuhlbusch, T. A. J.; Querol, X.; Alastuey, A.; Harrison, R. M.; Hopke, P. K.; Winiwarter, W.; Vallius, A.; Szidat, S.; Prevot, A. S. H.; Hueglin, C.; Bloemen, H.; Wahlin, P.; Vecchi, R.; Miranda, A. I.; Kasper-Giebl, A.; Maenhaut, W.; Hitzenberger, R., Source apportionment of particulate matter in Europe: A review of methods and results. *J Aerosol Sci* 2008, *39*, (10), 827-849.

20. Hu, Y. T.; Balachandran, S.; Pachon, J. E.; Jaemeen, B.; Ivey, C.; Holmes, H. A.; Odman, M. T.; Mulholland, J. A.; Russell, A. G., Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. In Atmospheric Chemistry and Physics 2014.

21. Ivey, C. E.; Holmes, H. A.; Hu, Y. T.; Mulholland, J. A.; Russell, A. G., Development of PM2.5 source impact spatial fields using a hybrid source apportionment air quality model. *Geosci Model Dev* 2015, *8*, (7), 2153-2165.

22. Darrow, L. A.; Klein, M.; Strickland, M. J.; Mulholland, J. A.; Tolbert, P. E., Ambient Air Pollution and Birth Weight in Full-Term Infants in Atlanta, 1994-2004. *Environ Health Persp* 2011, *119*, (5), 731-737.

23. Wang, X. B.; Ding, H.; Ryan, L.; Xu, X. P., Association between air pollution and low birth weight: A community-based study. *Environ Health Persp* 1997, *105*, (5), 514-520.

24. Campen, M. J.; Nolan, J. P.; Schladweiler, M. C. J.; Kodavanti, U. P.; Evansky, P. A.; Costa, D. L.; Watkinson, W. P., Cardiovascular and thermoregulatory effects of inhaled PM-associated transition metals: A potential interaction between nickel and vanadium sulfate. *Toxicol Sci* 2001, *64*, (2), 243-252.

25. Peel, J. L.; Klein, M.; Flanders, W. D.; Mulholland, J. A.; Freed, G.; Tolbert, P. E., Ambient Air Pollution and Apnea and Bradycardia in High-Risk Infants on Home Monitors. *Environ Health Persp* 2011, *119*, (9), 1321-1327.

26. Peters, A.; Wichmann, H. E.; Tuch, T.; Heinrich, J.; Heyder, J., Respiratory effects are associated with the number of ultrafine particles. *American Journal of Respiratory and Critical Care Medicine* 1997, *155*, (4), 1376-1383.

27. Pope, C. A.; Burnett, R. T.; Thun, M. J.; Calle, E. E.; Krewski, D.; Ito, K.; Thurston, G. D., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama-J Am Med Assoc* 2002, *287*, (9), 1132-1141.

28. Oberdorster, G.; Oberdorster, E.; Oberdorster, J., Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. *Environ Health Persp* 2005, *113*, (7), 823-839.

29. Dominici, F.; Peng, R. D.; Barr, C. D.; Bell, M. L., Protecting Human Health From Air Pollution Shifting From a Single-pollutant to a Multipollutant Approach. *Epidemiology* 2010, *21*, (2), 187-194.

30. Samet, J. M.; Dominici, F.; Curriero, F. C.; Coursac, I.; Zeger, S. L., Fine particulate air pollution and mortality in 20 US Cities, 1987-1994. *New England Journal of Medicine* 2000, *343*, (24), 1742-1749.

31. Sarnat, J. A., Fine Particle Sources and Cardiorespiratory Morbidity: An Application of Chemical Mass Balance and Factor Analytical Source Apportionment Methods. *Epidemiology* 2008, *19*, (6), S44-S44.

32. Tolbert, P. E.; Klein, M.; Peel, J. L.; Sarnat, S. E.; Sarnat, J. A., Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta. *J Expo Sci Env Epid* 2007, *17*, S29-S35.

33. Laden, F.; Neas, L. M.; Dockery, D. W.; Schwartz, J., Association of fine particulate matter from different sources with daily mortality in six US cities. *Environ Health Persp* 2000, *108*, (10), 941-947.

34. Thurston, G. D.; Ito, K.; Mar, T.; Christensen, W. F.; Eatough, D. J.; Henry, R. C.; Kim, E.; Laden, F.; Lall, R.; Larson, T. V.; Liu, H.; Neas, L.; Pinto, J.; Stolzel,

M.; Suh, H.; Hopke, P. K., Workgroup report: Workshop on source apportionment of particulate matter health effects - Intercomparison of results and implications. *Environ Health Persp* 2005, *113*, (12), 1768-1774.

35. Sarnat, J. A.; Marmur, A.; Klein, M.; Kim, E.; Russell, A. G.; Sarnat, S. E.; Mulholland, J. A.; Hopke, P. K.; Tolbert, P. E., Fine particle sources and cardiorespiratory morbidity: An application of chemical mass balance and factor analytical source-apportionment methods. *Environ Health Persp* 2008, *116*, (4), 459-466.

36. Balachandran, S.; Pachon, J. E.; Hu, Y. T.; Lee, D.; Mulholland, J. A.; Russell, A. G., Ensemble-trained source apportionment of fine particulate matter and method uncertainty analysis. *Atmospheric Environment* 2012, *61*, 387-394.

37. Seigneur, C.; Pun, B.; Pai, P.; Louis, J. F.; Solomon, P.; Emery, C.; Morris, R.; Zahniser, M.; Worsnop, D.; Koutrakis, P.; White, W.; Tombach, I., Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility. *Journal of the Air & Waste Management Association* 2000, *50*, (4), 588-599.

38. Schauer, J. J.; Cass, G. R., Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. *Environ Sci Technol* 2000, *34*, (9), 1821-1832.

39. Watson, J. G.; Cooper, J. A.; Huntzicker, J. J., The Effective Variance Weighting for Least-Squares Calculations Applied to the Mass Balance Receptor Model. *Atmospheric Environment* 1984, *18*, (7), 1347-1355.

40. Watson, J. G.; Chow, J. C.; Fujita, E. M., Review of volatile organic compound source apportionment by chemical mass balance. *Atmospheric Environment* 2001, *35*, (9), 1567-1584.

41. Maykut, N. N.; Lewtas, J.; Kim, E.; Larson, T. V., Source apportionment of PM2.5 at an urban IMPROVE site in Seattle, Washington. *Environ Sci Technol* 2003, *37*, (22), 5135-5142.

42. Paatero, P.; Tapper, U., Positive Matrix Factorization - a Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values. *Environmetrics* 1994, *5*, (2), 111-126.

43. Henry, R. C., Duality in multivariate receptor models. *Chemometr Intell Lab* 2005, *77*, (1-2), 59-63.

44. Marmur, A.; Park, S. K.; Mulholland, J. A.; Tolbert, P. E.; Russell, A. G., Source apportionment of PM2.5 in the southeastern United States using receptor and emissions-based models: Conceptual differences and implications for time-series health studies. *Atmospheric Environment* 2006, *40*, (14), 2533-2551.

45. Koo, B.; Wilson, G. M.; Morris, R. E.; Dunker, A. M.; Yarwood, G., Comparison of Source Apportionment and Sensitivity Analysis in a Particulate Matter Air Quality Model. *Environ Sci Technol* 2009, *43*, (17), 6669-6675.

46. Byun, D.; Schere, K. L., Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Appl Mech Rev* 2006, *59*, (1-6), 51-77.

47. Dunker, A. M., Efficient Calculation of Sensitivity Coefficients for Complex Atmospheric Models. *Atmospheric Environment* 1981, *15*, (7), 1155-1161.

48. Dunker, A. M., The Decoupled Direct Method for Calculating Sensitivity Coefficients in Chemical-Kinetics. *J Chem Phys* 1984, *81*, (5), 2385-2393.

49. Napelenok, S. L.; Cohan, D. S.; Hu, Y. T.; Russell, A. G., Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). *Atmospheric Environment* 2006, *40*, (32), 6112-6121.

50. Zhang, W.; Capps, S. L.; Hu, Y.; Nenes, A.; Napelenok, S. L.; Russell, A. G., Development of the high-order decoupled direct method in three dimensions for particulate matter: enabling advanced sensitivity analysis in air quality models. *Geosci Model Dev* 2012, *5*, (2), 355-368.

51. Bell, M. L., The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study of ozone in the Northern Georgia Region of the United States. *Environ Int* 2006, *32*, (5), 586-593.

52. Hu, Y.; Balachandran, S.; Pachon, J. E.; Baek, J.; Ivey, C.; Holmes, H.; Odman, M. T.; Mulholland, J. A.; Russell, A. G., Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. *Atmos. Chem. Phys.* 2014, *14*, (11), 5415-5431.

53. Hansen, D. A.; Edgerton, E.; Hartsell, B.; Jansen, J.; Burge, H.; Koutrakis, P.; Rogers, C.; Suh, H.; Chow, J.; Zielinska, B.; McMurry, P.; Mulholland, J.; Russell, A.; Rasmussen, R., Air quality measurements for the aerosol research and inhalation epidemiology study. *Journal of the Air & Waste Management Association* 2006, *56*, (10), 1445-1458.

54. Hansen, D. A.; Edgerton, E. S.; Hartsell, B. E.; Jansen, J. J.; Kandasamy, N.; Hidy, G. M.; Blanchard, C. L., The southeastern aerosol research and characterization study: Part 1-overview. *Journal of the Air & Waste Management Association* 2003, *53*, (12), 1460-1471.

55. Chow, J. C.; Watson, J. G.; Pritchett, L. C.; Pierson, W. R.; Frazier, C. A.; Purcell, R. G., The Dri Thermal Optical Reflectance Carbon Analysis System - Description, Evaluation and Applications in United-States Air-Quality Studies. *Atmos Environ a-Gen* 1993, *27*, (8), 1185-1201.

56. Baker, K. R.; Simon, H.; Kelly, J. T., Challenges to Modeling "Cold Pool" Meteorology Associated with High Pollution Episodes. *Environ Sci Technol* 2011, *45*, (17), 7118-7119.

57. Hu, Y.; Odman, M. T.; Russell, A. G., Mass conservation in the Community Multiscale Air Quality model. *Atmospheric Environment* 2006, *40*, 1199-1204.

58. Carter, W. P. L. *Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment*; Contract No. 92-329 and 95-308; California Air Resources Board: 2000.

59. Binkowski, F. S.; Roselle, S. J., Models-3 Community Multi-scale Air Quality (CMAQ) model aerosol component: 1. Model description. *Journal of Geophysical Research* 2003, *108*, (D6), 4183, doi:10.1029/2001JD001409.

60. Pleim, J. E.; Xiu, A., Development and Testing of a Surface Flux and Planetary Boundary-Layer Model for Application in Mesoscale Models. *J Appl Meteorol* 1995, *34*, (1), 16-32.

61. Xiu, A. J.; Pleim, J. E., Development of a land surface model. Part I: Application in a mesoscale meteorological model. *J Appl Meteorol* 2001, *40*, (2), 192-209.

62. Fletcher, R., *Practical Methods of Optimization*. John Wiley and Sons: 1987.

63. Reff, A.; Bhave, P. V.; Simon, H.; Pace, T. G.; Pouliot, G. A.; Mobley, J. D.; Houyoux, M., Emissions Inventory of PM2.5 Trace Elements across the United States. *Environ Sci Technol* 2009, *43*, (15), 5790-5796.

64. Balachandran, S.; Chang, H. H.; Pachon, J. E.; Holmes, H. A.; Mulholland, J. A.; Russell, A. G., Bayesian-Based Ensemble Source Apportionment of PM2.5. *Environ Sci Technol* 2013, *47*, (23), 13511-13518.

65. Hanna, S. R.; Chang, J. C.; Fernau, M. E., Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment* 1998, *32*, (21), 3619-3628.

66. Hanna, S. R.; Lu, Z. G.; Frey, H. C.; Wheeler, N.; Vukovich, J.; Arunachalam, S.; Fernau, M.; Hansen, D. A., Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric Environment* 2001, *35*, (5), 891-903.

67. Gillies, R. R.; Wang, S. Y.; Booth, M. R., Atmospheric Scale Interaction on Wintertime Intermountain West Low-Level Inversions. *Weather Forecast* 2010, *25*, (4), 1196-1210.

68. Kelly, K. E.; Kotchenruther, R.; Kuprov, R.; Silcox, G. D., Receptor model source attributions for Utah's Salt Lake City airshed and the impacts of wintertime

secondary ammonium nitrate and ammonium chloride aerosol. *Journal of the Air & Waste Management Association* 2013, *63*, (5), 575-590.

69. Pancras, J. P.; Landis, M. S.; Norris, G. A.; Vedantham, R.; Dvonch, J. T., Source apportionment of ambient fine particulate matter in Dearborn, Michigan, using hourly resolved PM chemical composition data. *Science of the Total Environment* 2013, *448*, 2-13.

70. Clements, A. L.; Fraser, M. P.; Upadhyay, N.; Herckes, P.; Sundblom, M.; Lantz, J.; Solomon, P. A., Chemical characterization of coarse particulate matter in the Desert Southwest - Pinal County Arizona, USA. *Atmospheric Pollution Research* 2014, *5*, (1), 52-61.

71. Chen, J. J.; Lu, J.; Avise, J. C.; DaMassa, J. A.; Kleeman, M. J.; Kaduwela, A. P., Seasonal modeling of PM2.5\ in California's San Joaquin Valley. *Atmospheric Environment* 2014, *92*, 182-190.

72. Ward, T.; Trost, B.; Conner, J.; Flanagan, J.; Jayanty, R. K. M., Source Apportionment of PM2.5 in a Subarctic Airshed - Fairbanks, Alaska. *Aerosol Air Qual Res* 2012, *12*, (4), 536-543.

73. Vedantham, R.; Landis, M. S.; Olson, D.; Pancras, J. P., Source Identification of PM2.5 in Steubenville, Ohio Using a Hybrid Method for Highly Time-Resolved Data. *Environ Sci Technol* 2014, *48*, (3), 1718-1726.

74. Tunno, B. J.; Shields, K. N.; Lioy, P.; Chu, N. J.; Kadane, J. B.; Parmanto, B.; Pramana, G.; Zora, J.; Davidson, C.; Holguin, F.; Clougherty, J. E., Understanding intra-neighborhood patterns in PM2.5 and PM10 using mobile monitoring in Braddock, PA. *Environ Health-Glob* 2012, *11*.

75. Malm, W. C.; Schichtel, B. A.; Pitchford, M. L., Uncertainties in PM2.5 Gravimetric and Speciation Measurements and What We Can Learn from Them. *Journal of the Air & Waste Management Association* 2011, *61*, (11), 1131-1149.

76. Cohan, D. S.; Hakami, A.; Hu, Y. T.; Russell, A. G., Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis. *Environ Sci Technol* 2005, *39*, (17), 6739-6748.

77. CEP, Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) User Manual, edited. In The University of North Carolina at Chapel Hill: Chapel Hill, NC, 2003.

78. Chester, S. E., State Recommendation for 24-hr PM2.5 NAAQS. In State of Michigan Department of Environmental Quality: Lansing Michigan, 2007.

79. Brewer, J. K., Pinal County PM10 and PM2.5 Area Designation Recommendations. In State of Arizona Governor's Office: Phoenix, AZ, 2010.

80. Xie, Y. L.; Hopke, P. K.; Paatero, P., Positive matrix factorization applied to a curve resolution problem. *J Chemometr* 1998, *12*, (6), 357-364.

81. Associates, E. H. P., SPECIATE 4.2: speciation Database Development Documentation. In Springfield, Virginia, 2009.

82. Cass, G. R., Organic molecular tracers for particulate air pollution sources. *Trac-Trend Anal Chem* 1998, *17*, (6), 356-366.

83. Zheng, M.; Cass, G. R.; Schauer, J. J.; Edgerton, E. S., Source apportionment of PM2.5 in the southeastern United States using solvent-extractable organic compounds as tracers. *Environ Sci Technol* 2002, *36*, (11), 2361-2371.

84. Zheng, M.; Cass, G. R.; Ke, L.; Wang, F.; Schauer, J. J.; Edgerton, E. S.; Russell, A. G., Source apportionment of daily fine particulate matter at Jefferson street, Atlanta, GA, during summer and winter. *Journal of the Air & Waste Management Association* 2007, *57*, (2), 228-242.

85. Lee, S.; Russell, A. G., Estimating uncertainties and uncertainty contributors of CMB PM2.5 source apportionment results. *Atmospheric Environment* 2007, *41*, (40), 9616-9624.

86. Shi, G. L.; Tian, Y. Z.; Zhang, Y. F.; Ye, W. Y.; Li, X.; Tie, X. X.; Feng, Y. C.; Zhu, T., Estimation of the concentrations of primary and secondary organic carbon in ambient particulate matter: Application of the CMB-Iteration method. *Atmospheric Environment* 2011, *45*, (32), 5692-5698.

87. Maier, M. L.; Balachandran, S.; Sarnat, S. E.; Turner, J. R.; Mulholland, J. A.; Russell, A. G., Application of an Ensemble-Trained Source Apportionment Approach at a Site Impacted by Multiple Point Sources. *Environ Sci Technol* 2013, *47*, (8), 3743-3751.

88. Nguyen, H. V.; Bai, L. In *Cosine similarity metric learning for face verification*, Asian Conference on Computer Vision, 2010; Springer: 2010; pp 709-720.

89. Binkowski, F. S.; Roselle, S. J., Models-3 community multiscale air quality (CMAQ) model aerosol component - 1. Model description. *J. Geophys. Res.-Atmos.* 2003, *108*, (D6).

90. Nenes, A.; Pandis, S. N.; Pilinis, C., ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. *Aquatic Geochem.* 1998, *4*, (1), 123-152.

91. Foley, K. M.; Roselle, S. J.; Appel, K. W.; Bhave, P. V.; Pleim, J. E.; Otte, T. L.; Mathur, R.; Sarwar, G.; Young, J. O.; Gilliam, R. C.; Nolte, C. G.; Kelly, J. T.; Gilliland, A. B.; Bash, J. O., Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. *Geosci Model Dev* 2010, *3*, (1), 205-226.

92. Kelly, J. T.; Bhave, P. V.; Nolte, C. G.; Shankar, U.; Foley, K. M., Simulating emission and chemical evolution of coarse sea-salt particles in the Community Multiscale Air Quality (CMAQ) model. *Geosci Model Dev* 2010, *3*, (1), 257-273.

93. Shrivastava, M.; Fast, J.; Easter, R.; Gustafson, W. I.; Zaveri, R. A.; Jimenez, J. L.; Saide, P.; Hodzic, A., Modeling organic aerosols in a megacity: comparison of simple and complex representations of the volatility basis set approach. *Atmospheric Chemistry and Physics* 2011, *11*, (13), 6639-6662.

94. Kim, S. W.; Heckel, A.; Frost, G. J.; Richter, A.; Gleason, J.; Burrows, J. P.; McKeen, S.; Hsie, E. Y.; Granier, C.; Trainer, M., NO2 columns in the western United States observed from space and simulated by a regional chemistry model and their implications for NOx emissions. *J. Geophys. Res.-Atmos.* 2009, *114*.

95. Fountoukis, C.; Nenes, A., ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+-Ca2+-Mg2+-Nh(4)(+)-Na+-SO42--NO3--Cl--H2O aerosols. *Atmospheric Chemistry and Physics* 2007, *7*, (17), 4639-4659.

96. Pandis, S. N.; Harley, R. A.; Cass, G. R.; Seinfeld, J. H., Secondary Organic Aerosol Formation and Transport. *Atmos Environ a-Gen* 1992, *26*, (13), 2269-2282.

97. Russell, A. G.; Mcrae, G. J.; Cass, G. R., Mathematical-Modeling of the Formation and Transport of Ammonium-Nitrate Aerosol. *Atmospheric Environment* 1983, *17*, (5), 949-964.

98. Hildemann, L. M.; Cass, G. R.; Mazurek, M. A.; Simonelt, B. R. T., Mathematical-Modeling of Urban Organic Aerosol - Properties Measured by High-Resolution Gas-Chromatography. *Environ Sci Technol* 1993, *27*, (10), 2045-2055.

99. Simon, H.; Baker, K. R.; Phillips, S., Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. *Atmospheric Environment* 2012, *61*, 124-139.

100. Kanakidou, M.; Seinfeld, J. H.; Pandis, S. N.; Barnes, I.; Dentener, F. J.; Facchini, M. C.; Van Dingenen, R.; Ervens, B.; Nenes, A.; Nielsen, C. J.; Swietlicki, E.; Putaud, J. P.; Balkanski, Y.; Fuzzi, S.; Horth, J.; Moortgat, G. K.; Winterhalter, R.; Myhre, C. E. L.; Tsigaridis, K.; Vignati, E.; Stephanou, E. G.; Wilson, J., Organic aerosol and global climate modelling: a review. *Atmospheric Chemistry and Physics* 2005, *5*, 1053-1123.

101. Bessagnet, B.; Hodzic, A.; Vautard, R.; Beekmann, M.; Cheinet, S.; Honore, C.; Liousse, C.; Rouil, L., Aerosol modeling with CHIMERE - preliminary evaluation at the continental scale. *Atmospheric Environment* 2004, *38*, (18), 2803-2817.

102. Geocounts Georgia Department of Transportation: Traffic Counts in Georgia. http://geocounts.com/gdot/

103. Administration, U. S. E. I. California: State Profile and Energy Estimates. http://www.eia.gov/state/?sid=PA

104. Administration, U. S. E. I. Pennsylvania: State Profile and Energy Estimates. http://www.eia.gov/state/?sid=PA

105. National Emissions Inventory: 2005-Based Modeling Platform. In U.S. Evironmental Protection Agency: 2011.

APPENDIX A

CHAPTER 2 SUPPLEMENTAL INFORMATION

Note A

During the chosen period of study, the CSN network performed thermal optical transmittance analyses to determine carbon species concentrations. As a result, organic and elemental carbon concentrations were converted to IMPROVE or total optical reflectance equivalents using previously derived methods (Malm et al. 2011). Additionally, measurements below detection limit are handled by setting the concentration to ½ of the measurement detection limit (MDL) and setting the measurement uncertainty to ⅔ of the MDL (Marmur et al. 2006).

Table A.1. Median and mean CTM–RM and SH adjustment factors for withheld CSN observation locations

Source Categories	$\widetilde{R_{CTM-RM}}$	$\widetilde{R_{SH}}$	$\overline{R_{CTM-RM}}$	$\overline{R_{SH}}$	**RMSE**
Agricultural Burning	0.881	0.778	0.669	0.683	0.316
Aircraft Emissions	1.000	1.000	0.996	0.997	0.022
Biogenic Emissions	1.001	1.002	0.986	0.988	0.126
Coal CM[1]	0.954	0.966	0.937	0.941	0.074
Diesel CM[1]	1.000	1.000	1.000	1.000	0.004
Dust	0.100	0.105	0.131	0.133	0.185
Fuel Oil CM[1]	0.985	0.969	0.956	0.956	0.100
Livestock Emissions	1.001	0.999	0.990	0.994	0.094
Liquid Petroleum Gas CM[1]	1.000	1.000	0.999	0.999	0.009
Lawn Waste Burning	0.100	0.107	0.285	0.156	1.738
Metal Processing	0.778	0.742	0.677	0.687	0.322
Meat Cooking	0.985	0.973	0.931	0.937	0.168
Mexican CM[1]	1.000	1.000	0.999	1.001	0.017
Mineral Processing	0.875	0.874	0.857	0.845	0.165
Natural Gas CM[1]	0.604	0.620	0.599	0.591	0.269
NR[s] Diesel CM[1]	1.001	0.995	0.964	0.979	0.132
NR[s] Fuel Oil CM[1]	0.999	0.996	0.985	0.990	0.055
NR[s] Gasoline CM[1]	1.000	0.998	0.963	0.975	0.139
NR[s] Liquid Petroleum Gas CM[1]	1.000	1.000	0.997	0.998	0.017
NR[s] Natural Gas CM[1]	1.000	1.000	0.999	1.000	0.007
Other NR Sources	1.000	1.000	1.000	0.999	0.005
Open Fires	0.704	0.652	0.584	0.590	0.270
Onroad Diesel CM[1]	0.983	0.976	0.944	0.960	0.115
Onroad Gasoline CM[1]	0.873	0.884	0.826	0.857	0.285
Other CM[1] Sources	0.954	0.911	0.874	0.867	0.254
Other PM Sources	0.667	0.654	0.632	0.641	0.279
Prescribed Burning	0.959	0.942	1.047	0.894	1.930
Railroad Emissions	1.001	0.998	0.992	0.993	0.046
Seasalt	0.995	0.993	0.988	0.986	0.034
Solvent Emissions	0.943	0.920	0.843	0.852	0.262
Wildfires	0.866	0.866	0.791	0.787	0.250
Woodfuel Burning	0.953	0.939	0.854	0.881	0.264
Wood stoves	0.100	0.128	0.322	0.184	1.768

Table A.2. January 2004 domain-wide $PM_{2.5}$ emissions and uncertainties.

Source Categories	Emissions (metric tons/day)	Uncertainty (factor)	Rank
Agricultural Burning	323.5	± 5	12
Aircraft Emissions	25.1	± 1.5	25
Biogenic Emissions	0.0	± 1.5	33
Coal CM[1]	1274.3	± 1.1	4
Diesel CM[1]	2.0	± 1.3	30
Dust	5300.9	± 10	1
Fuel Oil CM[1]	228.6	± 1.3	17
Livestock Emissions	1.3	± 1.3	31
Liquid Petroleum Gas CM[1]	13.7	± 1.3	26
Lawn Waste Burning	759.9	± 5	6
Metal Processing	171.2	± 1.3	18
Meat Cooking	309.5	± 1.5	14
Mexican CM[1]	13.2	± 1.5	27
Mineral Processing	315.9	± 1.3	13
Natural Gas CM[1]	619.7	± 1.3	8
NR[§] Diesel CM[1]	353.7	± 1.5	11
NR[§] Fuel Oil CM[1]	33.8	± 1.5	24
NR[§] Gasoline CM[1]	238.3	± 1.5	16
NR[§] Liquid Petroleum Gas CM[1]	2.9	± 1.5	29
NR[§] Natural Gas CM[1]	0.3	± 1.5	32
Other NR Sources	5.8	± 1.5	28
Open Fires	398.5	± 5	9
Onroad Diesel CM[1]	358.3	± 1.5	10
Onroad Gasoline CM[1]	159.7	± 1.5	20
Other CM[1] Sources	169.6	± 1.5	19
Other PM Sources	785.0	± 1.5	5
Prescribed Burning	725.3	± 10	7
Railroad Emissions	64.4	± 1.5	21
Seasalt	1893.4	± 1.5	3
Solvent Emissions	61.3	± 1.5	22
Wildfires	45.7	± 10	23
Woodfuel Burning	245.8	± 1.5	15
Wood stoves	3407.4	± 5	2

Table A.3. Mean observations and model simulations for January 2004 for withheld CSN observations (N = 75 observations).

	OBS		**CMAQ–DDM**			**HYB**			**SH**		
	Avg.	Std.	Avg.	Std.	RMSE	Avg.	Std.	RMSE	Avg.	Std.	RMSE
$PM_{2.5}$	11.7	8.3	16.3	11	12.5	8.59	4.7	7.9	9.20	5.7	8.6
OC	2.05	2.1	3.28	3.3	3.6	1.23	0.84	1.9	1.39	1.1	2.1
EC	0.727	0.64	0.944	1.3	1.3	0.55	0.59	0.74	0.627	0.90	0.99
NO_3	2.96	3.8	2.05	2.3	3.6	1.87	2.0	3.5	1.87	2.1	3.6
NH_4	1.38	1.3	1.45	0.99	1.2	1.20	0.79	1.1	1.24	0.83	1.2
SO_4	2.07	1.4	2.78	1.9	1.5	2.32	1.6	1.1	2.38	1.7	1.2
Na	0.0602	0.069	0.0960	0.079	0.12	0.0339	0.019	0.074	0.0394	0.028	0.078
Mg	0.0127	0.0017	0.0255	0.020	0.028	0.0960	0.061	0.018	0.0111	0.091	0.019
Al	0.0149	0.015	0.187	0.16	0.23	0.0397	0.022	0.034	0.0451	0.029	0.043
Si	0.0814	0.066	0.568	0.50	0.69	0.113	0.070	0.081	0.149	0.017	0.18
P	0.0046	0.0019	0.0064	0.0054	0.0061	0.0021	0.0011	0.0034	0.0026	0.0018	0.0035
Cl	0.0436	0.10	0.369	0.33	0.46	0.0798	0.056	0.090	0.0931	0.077	0.12
K	0.0725	0.068	0.429	0.46	0.59	0.0738	0.062	0.065	0.0946	0.13	0.15
Ca	0.0407	0.060	0.182	0.14	0.19	0.0445	0.023	0.054	0.0501	0.030	0.058
V	0.022	0.0030	0.0015	0.0014	0.025	4.64 E-4	3.1E-4	0.0068	5.68E-4	5.0E-4	0.011
Fe	0.0707	0.16	0.160	0.14	0.0034	0.0368	0.022	0.0034	0.0447	0.036	0.0034
Cu	0.0027	0.0034	0.0035	0.053	0.0060	0.0015	0.0011	0.0032	0.0020	0.0022	0.0035
Zn	0.0209	0.070	0.0099	0.013	0.010	0.0029	0.0021	0.0075	0.0035	0.0032	0.0077
Se	0.0018	0.0016	0.0017	0.011	0.21	0.0012	7.1E-4	0.15	0.0012	7.7E-4	0.15
Pb	0.0054	0.011	0.0019	0.0022	0.0064	5.60E-4	3.8E-4	0.0035	7.42E-4	7.8E-4	0.0038

Table A.4. Linear regression and correlation coefficients for model simulations vs. observations for January 2004 for withheld CSN observations (N = 75 observations). Regression equation: 〖Conc〗_model=α+ β〖·Conc〗_obs

	CMAQ–DDM vs. OBS					HYB vs. OBS					SH vs. OBS				
	α	SE_α	β	SE_β	r	α	SE_α	β	SE_β	r	α	SE_α	β	SE_β	r
$PM_{2.5}$	11.2	4.3	0.43	0.30	0.321	5.5	1.7	0.27	0.12	0.470	6.4	2.1	0.24	0.15	0.350
OC	2.49	1.0	0.39	0.36	0.242	0.73	0.22	0.25	0.076	0.605	1.01	0.33	0.18	0.11	0.351
EC	0.61	0.43	0.46	0.45	0.234	0.34	0.20	0.29	0.20	0.314	0.41	0.31	0.30	0.32	0.217
NO_3	1.3	0.63	0.25	0.13	0.409	1.1	0.54	0.25	0.11	0.454	1.2	0.55	0.25	0.12	0.440
NH_4	0.98	0.30	0.34	0.16	0.454	0.78	0.23	0.31	0.12	0.516	0.81	0.24	0.31	0.13	0.492
SO_4	0.77	0.56	0.97	0.23	0.706	0.48	0.46	0.89	0.19	0.744	0.55	0.48	0.89	0.19	0.730
Na	0.11	0.024	-0.17	0.27	-0.145	0.032	0.0059	0.030	0.065	0.110	0.040	0.0087	-0.016	0.096	-0.039
Mg	0.025	0.0057	0.038	0.27	0.033	0.0091	0.0018	0.039	0.084	0.109	0.011	0.0027	0.025	0.13	0.046
Al	0.16	0.050	1.86	2.4	0.182	0.033	0.0070	0.44	0.33	0.301	0.041	0.0093	0.27	0.44	0.145
Si	0.37	0.18	2.46	1.7	0.320	0.079	0.024	0.42	0.23	0.392	0.13	0.061	0.20	0.59	0.078
P	0.0081	0.0033	-0.36	0.67	-0.124	0.0024	6.90E-4	-0.058	0.14	-0.098	0.0035	0.011	-0.21	0.22	-0.216
Cl	0.35	0.081	0.46	0.72	0.147	0.065	0.011	0.33	0.098	0.620	0.086	0.019	0.16	0.17	0.221
K	0.42	0.16	0.11	1.6	0.016	0.041	0.019	0.45	0.19	0.496	0.098	0.043	-0.041	0.43	-0.022
Ca	0.13	0.0035	1.22	0.48	0.513	0.038	0.0060	0.16	0.083	0.420	0.044	0.008	0.16	0.11	0.316
V	0.0015	4.0E-4	-3.77E-4	0.11	-8.0E-4	4.25E-4	8.8E-5	0.018	0.024	0.173	5.45E-4	1.44E-4	0.011	0.039	0.063
Fe	0.15	0.035	0.16	0.20	0.180	0.034	0.0053	0.046	0.031	0.326	0.042	0.0090	0.038	0.053	0.166
Cu	0.0036	0.0016	-0.022	0.37	-0.014	0.0013	3.2E-4	0.078	0.074	0.238	0.0018	6.4E-4	0.071	0.15	0.110
Zn	0.0099	0.0033	4.8E-4	0.045	0.003	0.0029	5.2E-4	0.0011	0.0072	0.037	0.0034	7.4E-4	0.0032	0.011	0.070
Se	0.0015	3.9E-4	0.082	0.16	0.116	9.8E-4	2.5E-4	0.11	0.10	0.235	0.0010	2.7E-4	0.11	0.11	0.225
Pb	0.0017	5.7E-4	0.027	0.046	0.135	5.3E-4	9.7E-5	0.0052	0.0078	0.153	6.13E-4	7.9E-4	0.024	0.015	0.345

Table A.5. Mean observations and model simulations for January 2004 for SEARCH monitors (N = 8 monitors). All averages and standard deviations are expressed in μg m^{-3}.

	OBS		CMAQ–DDM			SH		
	Avg	Std.	Avg	Std.	RMSE	Avg	Std.	RMSE
$PM_{2.5}$	11.2	5.3	14.7	9.8	9.3	9.04	5.3	5.6
OC	2.81	2.4	3.05	2.8	1.8	1.66	1.5	1.3
EC	0.831	0.81	0.711	0.76	1.4	0.457	0.48	1.4
NO_3	1.08	0.78	1.01	1.5	0.93	0.927	1.4	0.75
NH_4	1.12	0.51	1.22	0.92	0.80	1.01	0.70	0.74
SO_4	2.73	1.6	3.78	2.3	2.8	3.30	2.0	2.3
Al	0.0062	0.0058	0.143	0.12	0.0015	0.0342	0.027	0.0016
Si	0.0292	0.026	0.388	0.28	0.010	0.0936	0.064	0.011
K	0.0600	0.027	0.279	0.24	0.0032	0.0504	0.035	0.0031
Ca	0.0281	0.031	0.141	0.11	0.0070	0.0396	0.032	0.0019
Fe	0.0368	0.044	0.132	0.14	0.0073	0.0328	0.031	0.0024
Cu	0.0013	0.0018	0.0043	0.0077	0.0041	0.0016	0.0017	0.0042
Zn	0.0157	0.030	0.0108	0.013	0.0020	0.0036	0.0028	0.0021
Se	8.00E-04	8.00E-04	0.0013	0.0013	0.0015	9.00E-04	9.00E-04	0.0012
Pb	0.0027	0.0041	0.0020	0.0033	0.0027	5.00E-04	7.00E-04	0.0029

Table A.6. Mean observations and model simulations for January 2004 for IMPROVE monitors (N = 38 monitors). All averages and standard deviations are expressed in µg m^{-3}.

	OBS		CMAQ–DDM			SH		
	Avg	Std.	Avg	Std.	RMSE	Avg	Std.	RMSE
$PM_{2.5}$	6.25	3.8	11.51	8.2	8.5	6.63	4.4	3.6
OC	1.39	1.2	2.33	2.3	2.3	0.959	1.2	1.3
EC	0.298	0.22	0.532	0.52	0.51	0.283	0.30	0.28
NO_3	1.25	1.4	1.41	1.6	1.2	1.23	1.4	1.1
NH_4	0.894	0.40	0.984	0.66	0.61	0.870	0.57	0.55
SO_4	1.64	1.1	2.43	1.7	1.5	2.09	1.4	1.08
Al	0.0128	0.017	0.103	0.099	0.13	0.0268	0.022	0.028
Si	0.0629	0.039	0.297	0.27	0.35	0.0781	0.067	0.071
K	0.0416	0.027	0.291	0.30	0.39	0.0615	0.060	0.059
Ca	0.0218	0.015	0.101	0.094	0.13	0.0297	0.024	0.024
Fe	0.0198	0.021	0.0860	0.094	0.11	0.0234	0.021	0.022
Cu	7.00E-04	6.0E-04	0.0020	0.0043	0.0044	0.0011	0.0019	0.0019
Zn	0.00840	0.0078	0.0085	0.010	0.010	0.0035	0.0066	0.0103
Se	9.00E-04	0.001	0.0011	0.0013	0.0014	8.00E-04	8.0E-04	0.001
Pb	0.00190	0.0024	0.0011	0.0016	0.0027	4.00E-04	5.0E-04	0.0028

Table A.7. Linear regression and correlation coefficients for model predictions vs. observations for January 2004 for SEARCH monitors (N = 8 monitors). Regression equation: 〖Conc〗_model=α+ β〖·Conc〗_obs

	CMAQ–DDM vs. OBS					SH vs. OBS				
	α	SE_α	**β**	SE_β	**r**	**α**	SE_α	**β**	SE_β	**r**
$PM_{2.5}$	6.40	5.2	0.746	0.42	0.404	6.67	2.7	0.332	0.22	0.357
OC	1.72	0.93	0.473	0.25	0.413	1.18	0.47	0.264	0.13	0.445
EC	0.353	0.23	0.431	0.20	0.462	0.263	0.14	0.307	0.12	0.519
NO_3	0.222	0.59	0.730	0.44	0.371	0.497	0.59	0.604	0.44	0.315
NH_4	0.760	0.53	0.411	0.43	0.227	0.827	0.40	0.294	0.32	0.216
SO_4	0.794	0.74	1.10	0.23	0.752	0.840	0.55	0.978	0.17	0.807
Al	0.0882	0.037	8.86	4.4	0.446	0.028	0.0097	1.89	1.1	0.379
Si	0.214	0.087	5.95	2.2	0.556	0.0755	0.022	1.12	0.58	0.437
K	0.135	0.14	2.40	2.1	0.271	0.0412	0.020	0.253	0.31	0.200
Ca	0.0820	0.031	2.10	0.74	0.581	0.0281	0.0095	0.613	0.23	0.559
Fe	0.0395	0.028	2.52	0.49	0.788	0.0206	0.0086	0.473	0.15	0.615
Cu	7.00E-04	0.0018	2.85	0.81	0.657	0.0012	5.00E-04	0.507	0.24	0.470
Zn	0.0088	0.0033	0.130	0.098	0.313	0.0035	7.00E-04	0.0296	0.022	0.323
Se	0.0012	4.00E-04	0.0946	0.39	0.060	0.001	3.00E-04	0.0657	0.28	0.060
Pb	0.0013	9.00E-04	0.265	0.19	0.335	5.00E-04	2.00E-04	0.0593	0.042	0.330

Table A.8. Linear regression and correlation coefficients for model predictions vs. observations for January 2004 for IMPROVE monitors (N = 38 monitors). Regression equation: 〖Conc〗_model=α+ β〖·Conc〗_obs

	CMAQ–DDM vs. OBS					SH vs. OBS				
	α	SE_α	**β**	SE_β	**r**	**α**	SE_α	**β**	SE_β	**r**
$PM_{2.5}$	3.53	1.4	1.28	0.19	0.590	1.96	0.72	0.747	0.098	0.637
OC	1.11	0.35	0.877	0.19	0.456	0.359	0.18	0.432	0.097	0.439
EC	0.189	0.084	1.15	0.23	0.479	0.0862	0.049	0.661	0.13	0.472
NO_3	0.431	0.17	0.782	0.092	0.684	0.346	0.15	0.712	0.081	0.696
NH_4	0.369	0.31	0.689	0.32	0.422	0.355	0.27	0.576	0.28	0.406
SO_4	0.617	0.25	1.11	0.13	0.692	0.451	0.20	0.997	0.10	0.728
Al	0.0894	0.013	1.05	0.62	0.182	0.023	0.0029	0.295	0.13	0.232
Si	0.180	0.054	1.86	0.73	0.268	0.0554	0.014	0.361	0.18	0.208
K	0.194	0.059	2.33	1.2	0.209	0.0274	0.011	0.820	0.22	0.370
Ca	0.0387	0.016	2.83	0.62	0.448	0.0155	0.0043	0.648	0.16	0.400
Fe	0.0455	0.012	2.05	0.42	0.466	0.0143	0.0028	0.461	0.095	0.469
Cu	3.00E-04	7.0E-04	2.42	0.78	0.319	5.00E-04	3.00E-04	0.842	0.35	0.256
Zn	0.0049	0.002	0.434	0.13	0.339	0.002	0.001	0.182	0.090	0.214
Se	7.00E-04	2.0E-04	0.403	0.13	0.321	5.00E-04	1.00E-04	0.330	0.083	0.399
Pb	8.00E-04	2.0E-04	0.143	0.075	0.204	3.00E-04	1.00E-04	0.0305	0.022	0.151

Table A.9. January 2004 domain-wide precursor emissions estimates (metric tons/day).

Source Categories	NOx	VOC	$PM_{2.5}$	SO_2	NH_3	CO	EC	POA
Agricultural	21.5	157.0	323.5	0.6	3.4	1826.6	12.9	216.7
Aircraft Emissions	282.7	74.1	25.1	26.0	0.1	801.9	16.5	7.3
Biogenic Emissions	2381.1	24552.5	0.0	0.0	0.0	6140.7	0.0	0.0
Coal CM[1]	14250.1	192.0	1274.3	35080.3	4.2	2000.6	12.7	254.9
Diesel CM[1]	32.9	2.6	2.0	3.2	0.0	10.4	1.5	0.4
Dust	3.5	1.1	5300.9	0.1	0.4	1.8	22.7	311.3
Fuel Oil CM[1]	7960.5	61.2	228.6	20482.6	23.5	485.1	23.8	18.5
Livestock	0.0	86.0	1.3	0.0	3244.8	0.0	0.0	0.1
LPG CM[1]	265.7	8.7	13.7	113.2	0.2	50.2	0.0	8.3
Lawn Waste Burn	173.7	657.1	759.9	15.4	0.4	4973.4	28.5	255.9
Metal Processing	0.3	34.1	171.2	0.3	0.0	42.2	2.1	166.2
Meat Cooking	370.9	224.4	309.5	3451.7	5.3	5279.1	2.1	40.0
Mexican CM[1]	133.9	5.1	13.2	595.3	9.1	21.4	0.0	0.0
Mineral Processing	917.4	88.1	315.9	812.3	6.1	539.0	4.7	24.0
Natural Gas CM[1]	10119.7	569.2	619.7	5760.1	72.5	2984.6	0.0	371.1
NR[s] Diesel CM[1]	5262.8	150.4	353.7	412.2	3.4	1841.1	262.1	79.4
NR[s] Fuel Oil CM[1]	727.8	20.5	33.8	428.0	0.1	86.3	2.9	3.6
NR[s] Gasoline CM[1]	614.1	7472.8	238.3	34.1	2.4	47448.5	19.1	187.3
NR[s] LPG CM[1]	639.6	159.6	2.9	0.6	0.2	2508.9	0.0	2.5
NR[s] Nat. Gas CM[1]	81.1	1.2	0.3	0.1	0.0	324.3	0.0	0.3
Other NR Sources	0.0	0.0	5.8	0.0	0.0	0.0	0.0	0.0
Open Fires	96.4	235.7	398.5	0.0	0.0	3258.3	15.9	267.0
Onroad Diesel CM[1]	9613.6	249.0	358.3	309.4	17.9	3732.2	229.0	98.8
Onroad Gas CM[1]	14749.9	15719.4	159.7	337.2	788.1	249971.1	21.4	57.0
Other CM[1] Sources	814.0	67.1	169.6	1155.5	16.6	871.8	12.0	56.2
Other PM Sources	1152.9	3057.4	785.0	3053.7	1427.3	4665.5	12.4	151.2
Prescribed Burning	170.6	229.0	725.3	42.6	30.6	7300.9	116.0	558.4
Railroad Emissions	2765.7	38.4	64.4	158.8	0.0	330.8	47.7	14.4
Seasalt	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Solvent Emissions	334.9	22842.9	61.3	403.9	4.8	86.2	2.2	32.7
Wildfires	12.1	20.1	45.7	3.0	5.1	547.7	7.3	35.2
Woodfuel Burning	343.9	48.3	245.8	194.4	2.0	1001.0	34.4	95.9
Wood stoves	326.6	9282.2	3407.4	50.0	32.9	25707.7	367.0	1927.2
All Sources	74619.7	86307.1	16414.5	72924.7	5701.6	374839.3	1277.1	5242.0

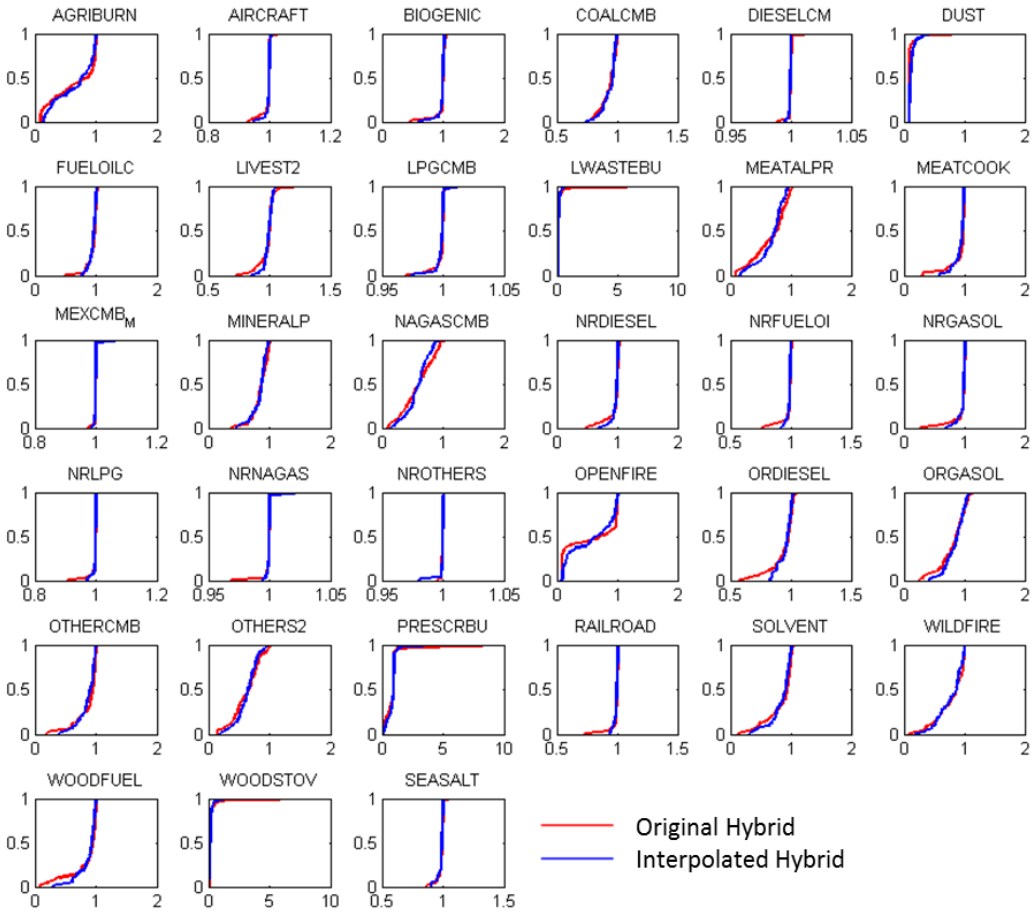

Figure A.1. Cumulative distributions of original and interpolated hybrid adjustment factors for witheld CSN observation locations.

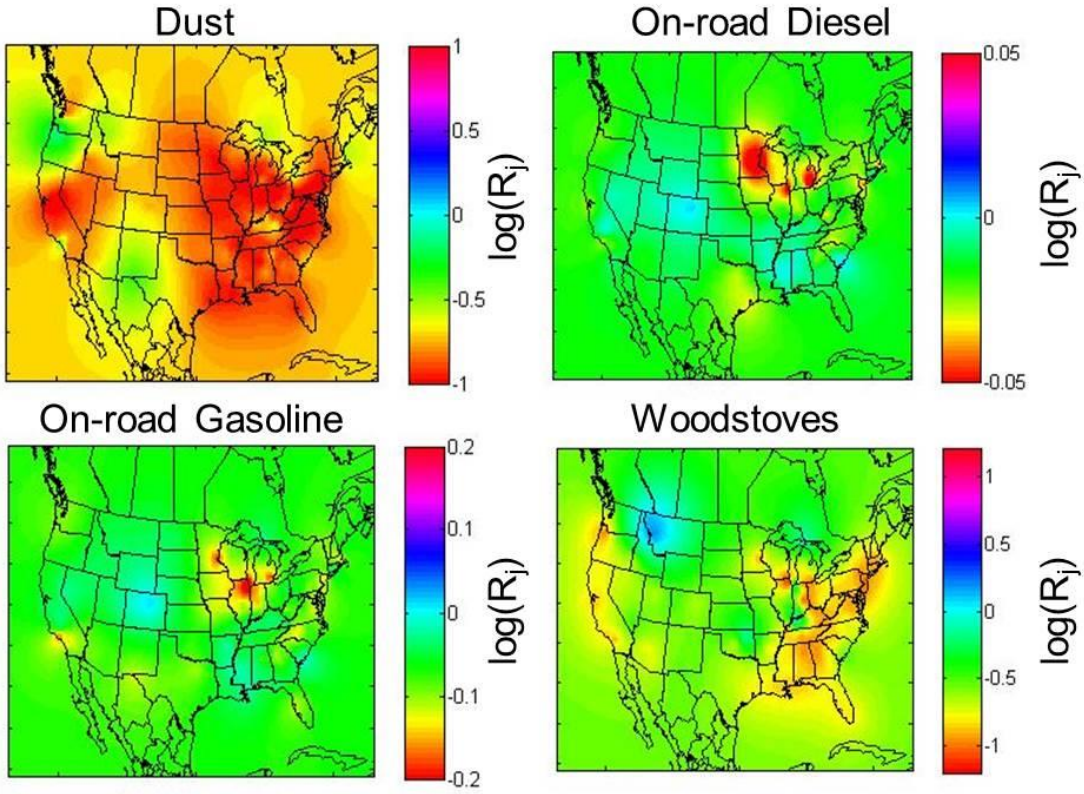

Figure A.2. Spatial fields of temporally-averaged, kriged adjustment factors (Rj_{SH}) for dust, on-road diesel combustion, on-road gasoline combustion, and wood stove sources for January 2004. Values were averaged over 9 observations days in the month. Note that each figure has a different scale.

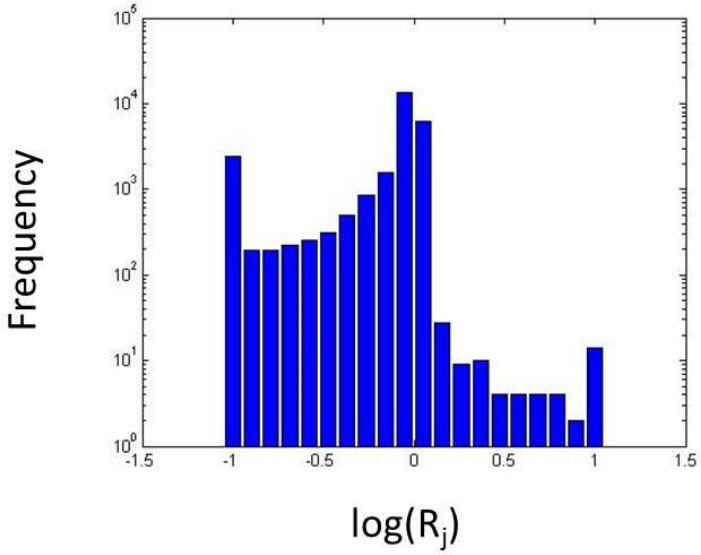

Figure A.3. Distribution of Rj values generated by hybrid analysis of CSN data for January 2004 (n = 26,400; 0.1 < Rj < 10). Note that the y-axis is on a log scale and Rj values are log transformed.

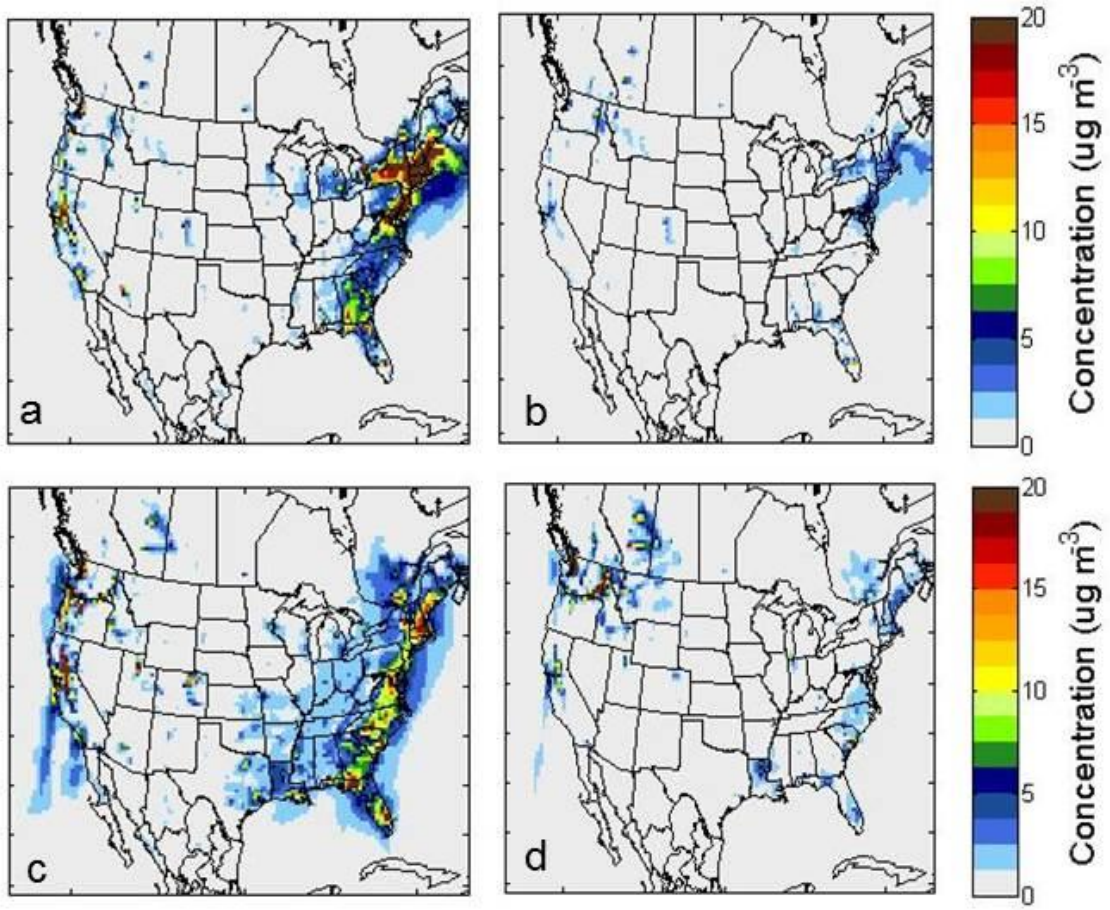

Figure A.4. Hybrid adjustment of biomass burning impacts on $PM_{2.5}$ on January 4th and 22nd in 2004. Biomass burning fields are produced by aggregating source impacts from agricultural burning, lawn waste burning, open fires, prescribed burning, wildfires, wood fuel and wood stove burning. (a) CMAQ–DDM spatial field for January 4th. (b) SH spatial field for January 4th. (c) CMAQ–DDM spatial field for January 22nd. (d) SH spatial field for January 22nd**.**

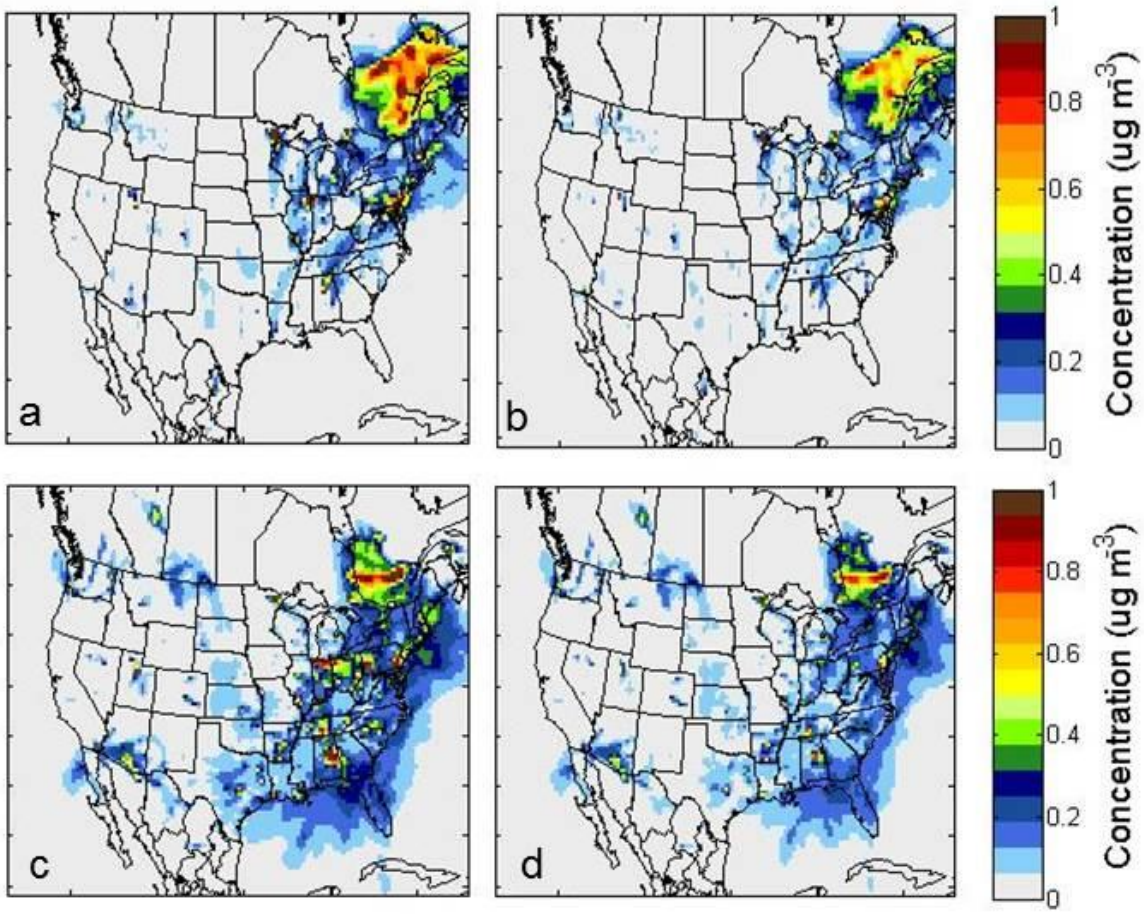

Figure A.5. Hybrid adjustment of metals processing impacts on $PM_{2.5}$ on January 4th and 22nd in 2004. (a) CMAQ–DDM spatial field for January 4th. (b) SH spatial field for January 4th. (c) CMAQ–DDM spatial field for January 22nd. (d) SH spatial field for January 22nd.

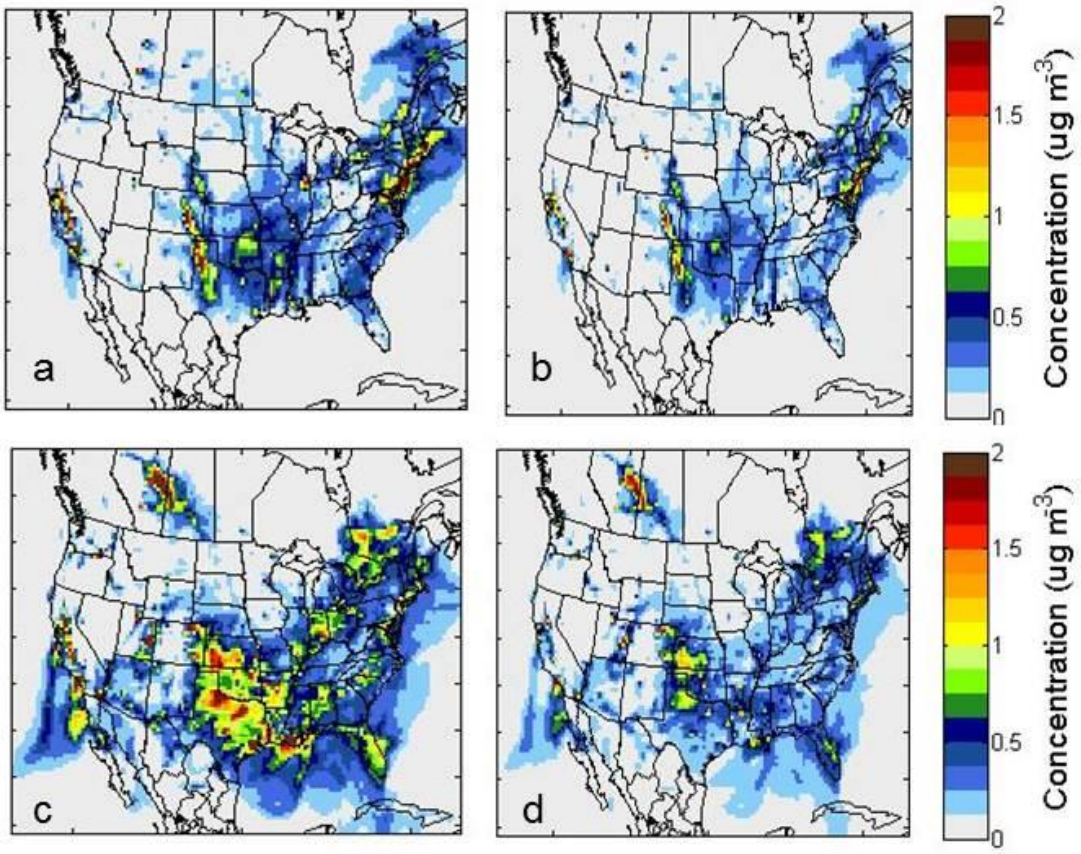

Figure A.6. Hybrid adjustment of natural gas combustion (point and area sources) source impact fields on January 4th and 22nd in 2004. (a) CMAQ–DDM spatial field for January 4th. (b) SH spatial field for January 4th. (c) CMAQ–DDM spatial field for January 22nd. (d) SH spatial field for January 22nd.

APPENDIX B

CHAPTER 3 SUPPLEMENTAL INFORMATION

Table B.1. Performance statistics for CMAQ-DDM and spatial hybrid (SH) $PM_{2.5}$ concentrations compared to observations at CSN locations for 2005-2007.

Year	**Observation**		**CMAQ-DDM**		**SH**	
concentration	mean	std.	mean	std.	mean	std.
2005	13.8	8.9	12.5	8.3	11.7	7.5
2006	12.3	7.6	12.4	8.4	10.2	6.5
2007	13.2	8.7	6.6	5.1	7.8	8.6
mean bias			v. observation		v. observation	
2005		-	-1.3		-2.1	
2006		-	0.0		-2.1	
2007		-	-6.7		-5.4	
			v. observation		v. observation	
2005		-	0.13		0.13	
2006		-	0.50		0.55	
2007		-	0.32		0.36	

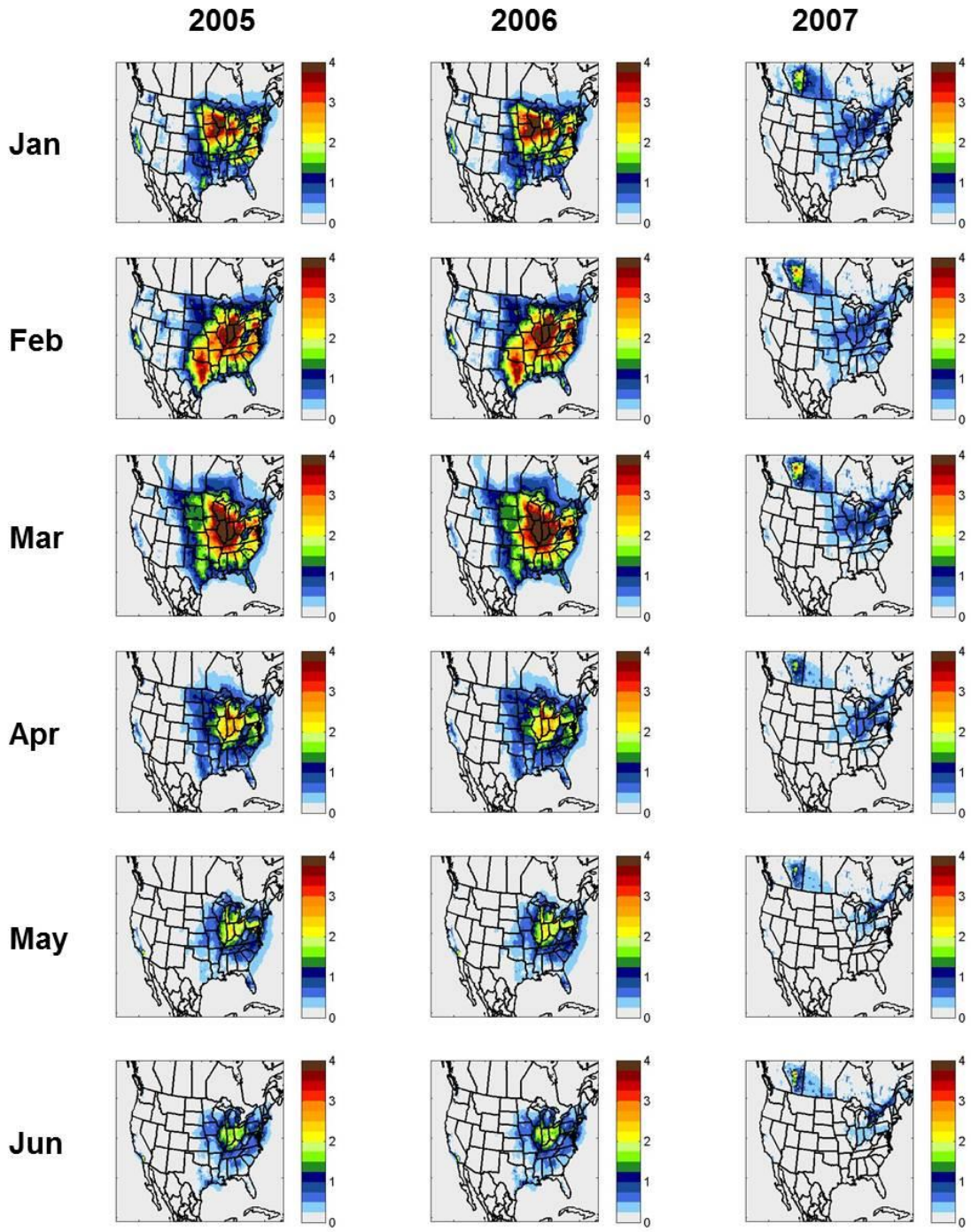

Figure B.1. Spatial fields of monthly-averaged optimized agricultural/livestock impacts on $PM_{2.5}$ for years 2005-2007.

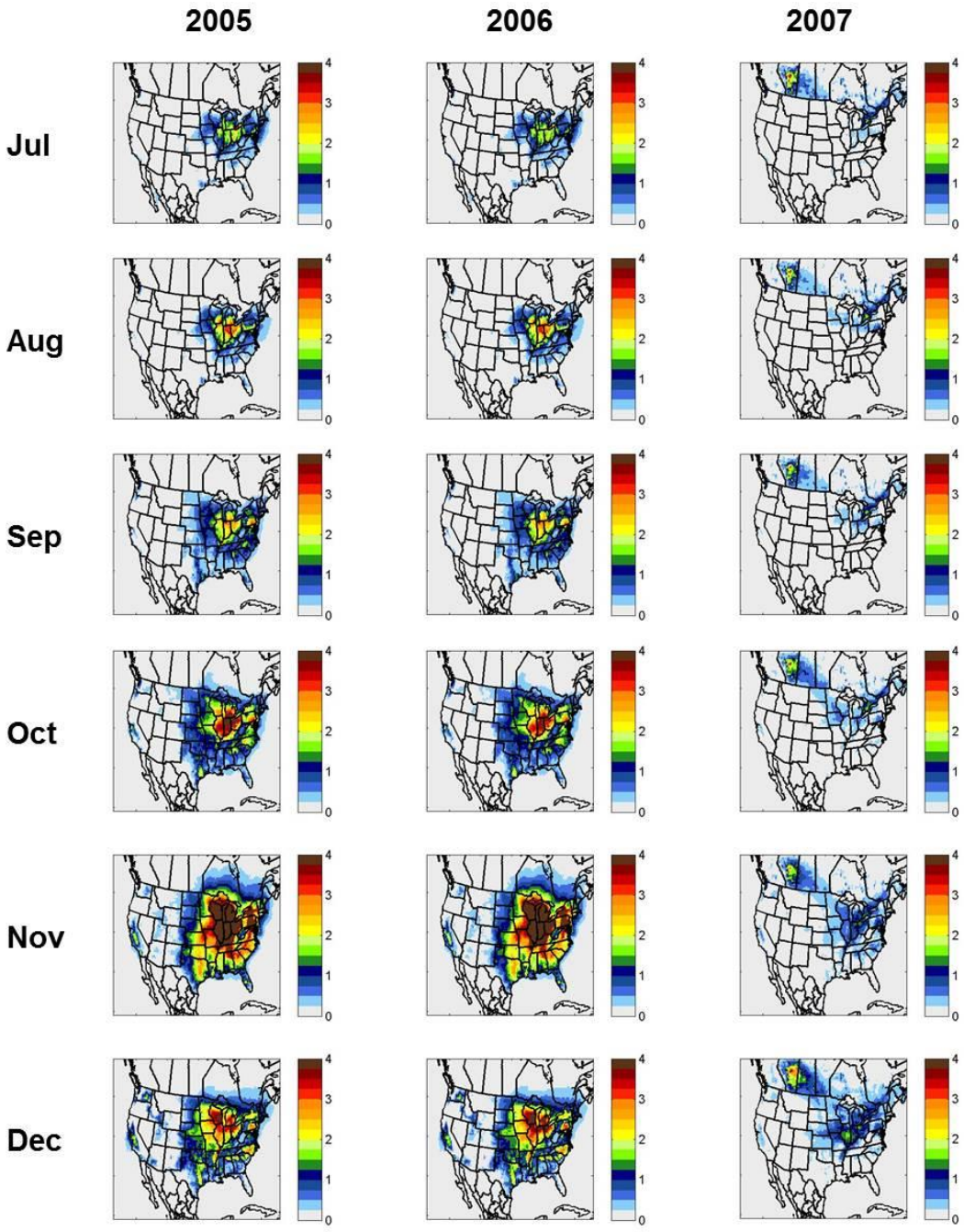

Figure B.1. Continued

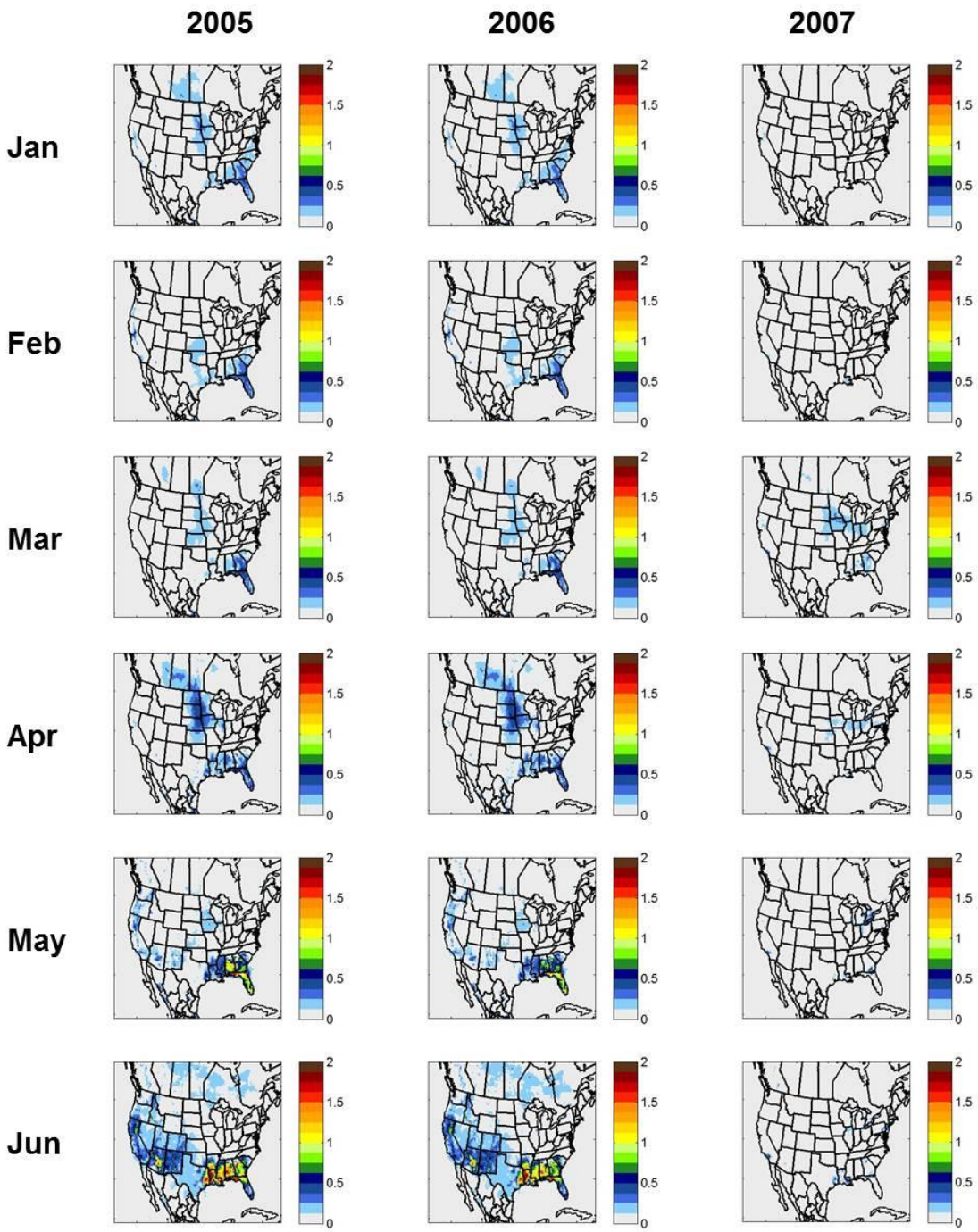

Figure B.2. Spatial fields of monthly-averaged, optimized biogenic impacts on $PM_{2.5}$ for years 2005-2007.

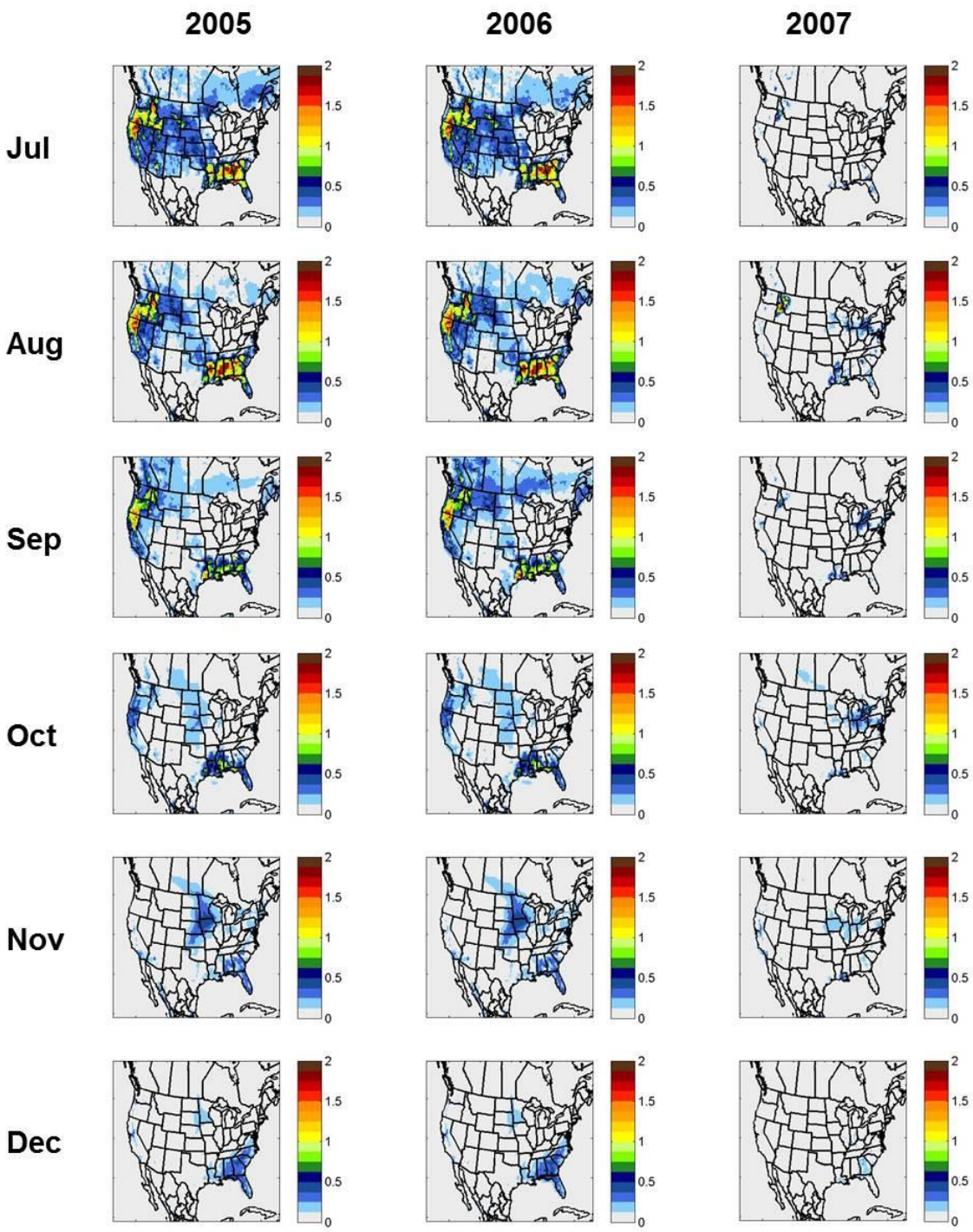

Figure B.2. Continued

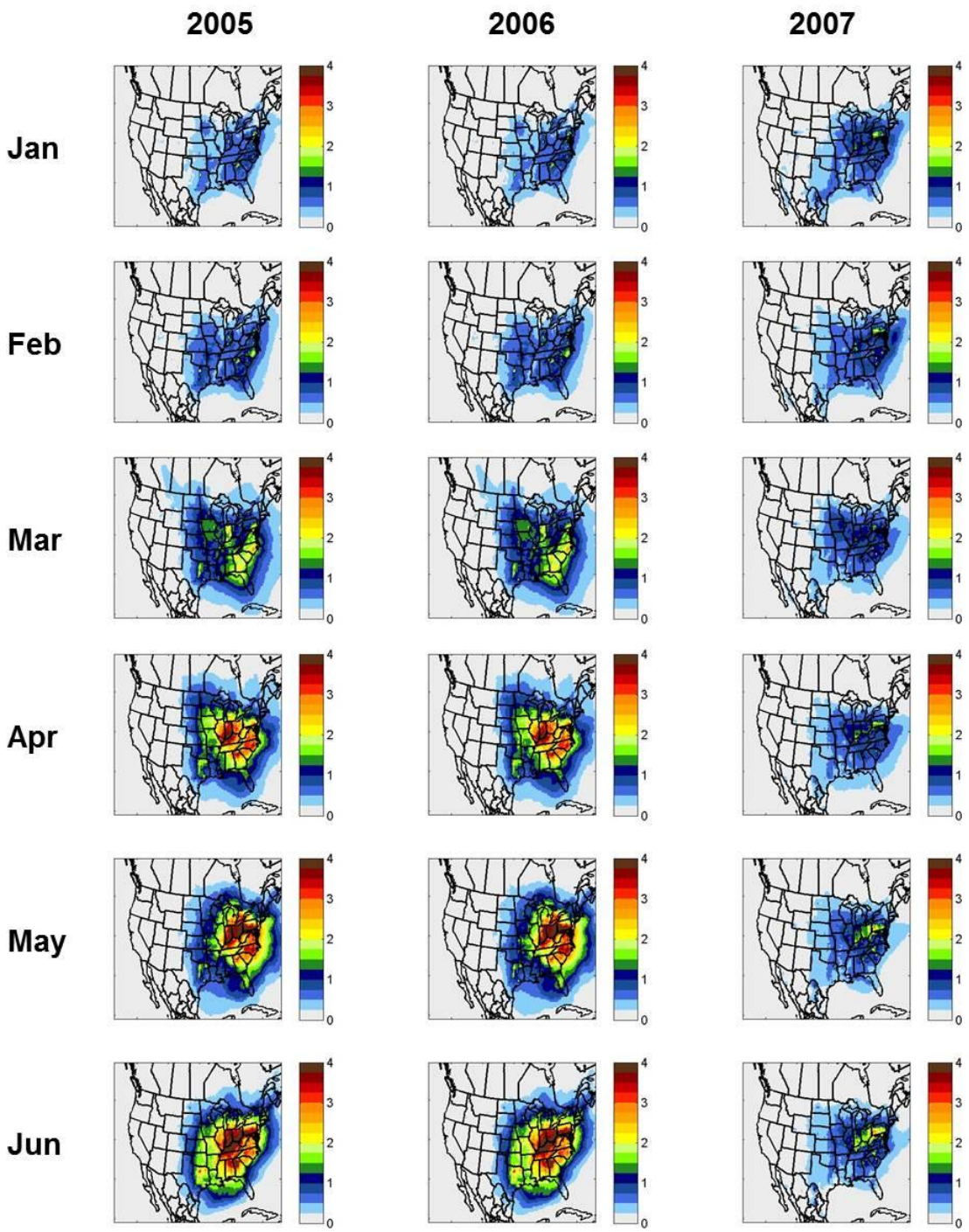

Figure B.3. Spatial fields of monthly-averaged optimized coal combustion impacts on $PM_{2.5}$ for years 2005-2007.

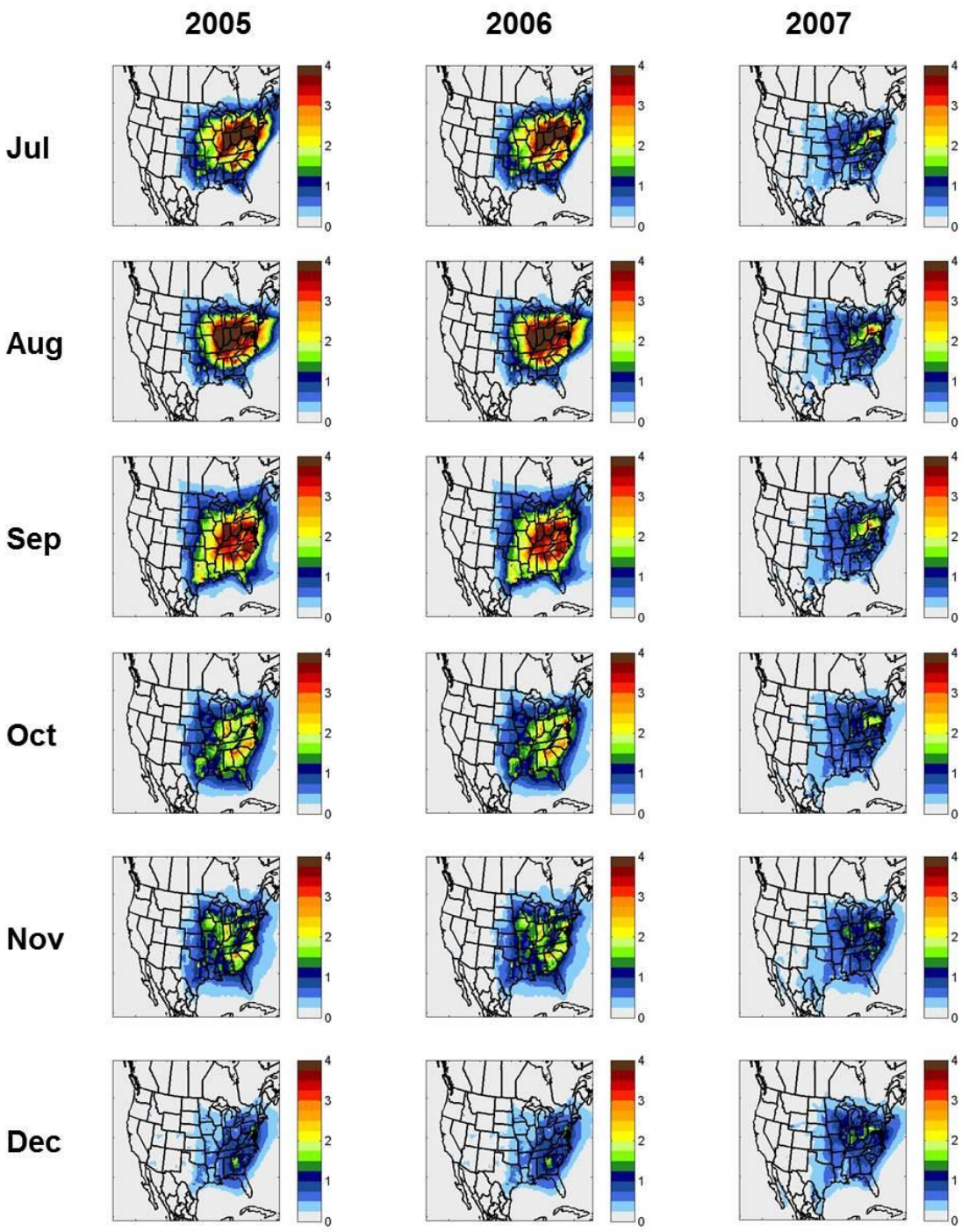

Figure B.3. Continued

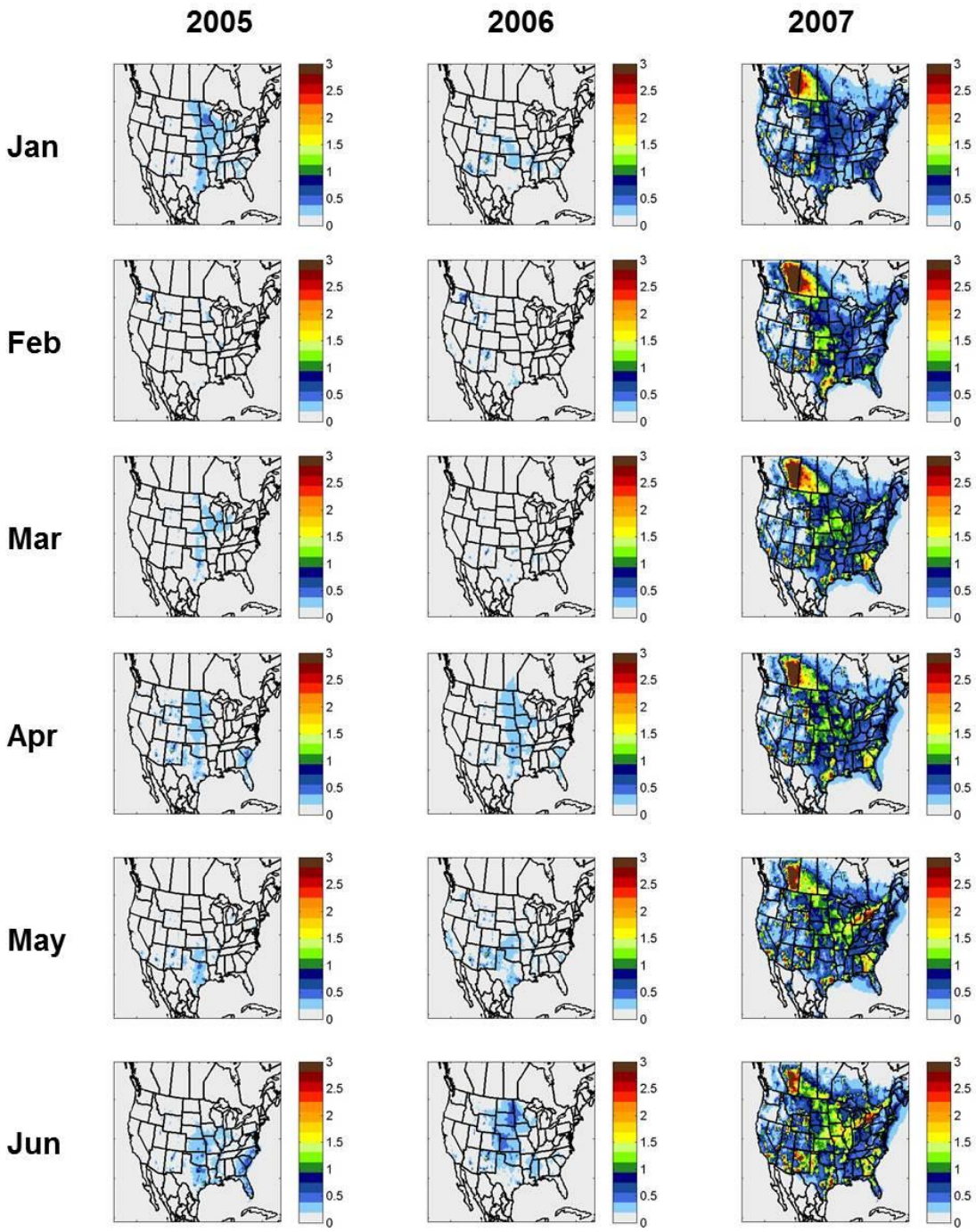

Figure B.4: Spatial fields of monthly-averaged dust impacts on $PM_{2.5}$ for years 2005-2007.

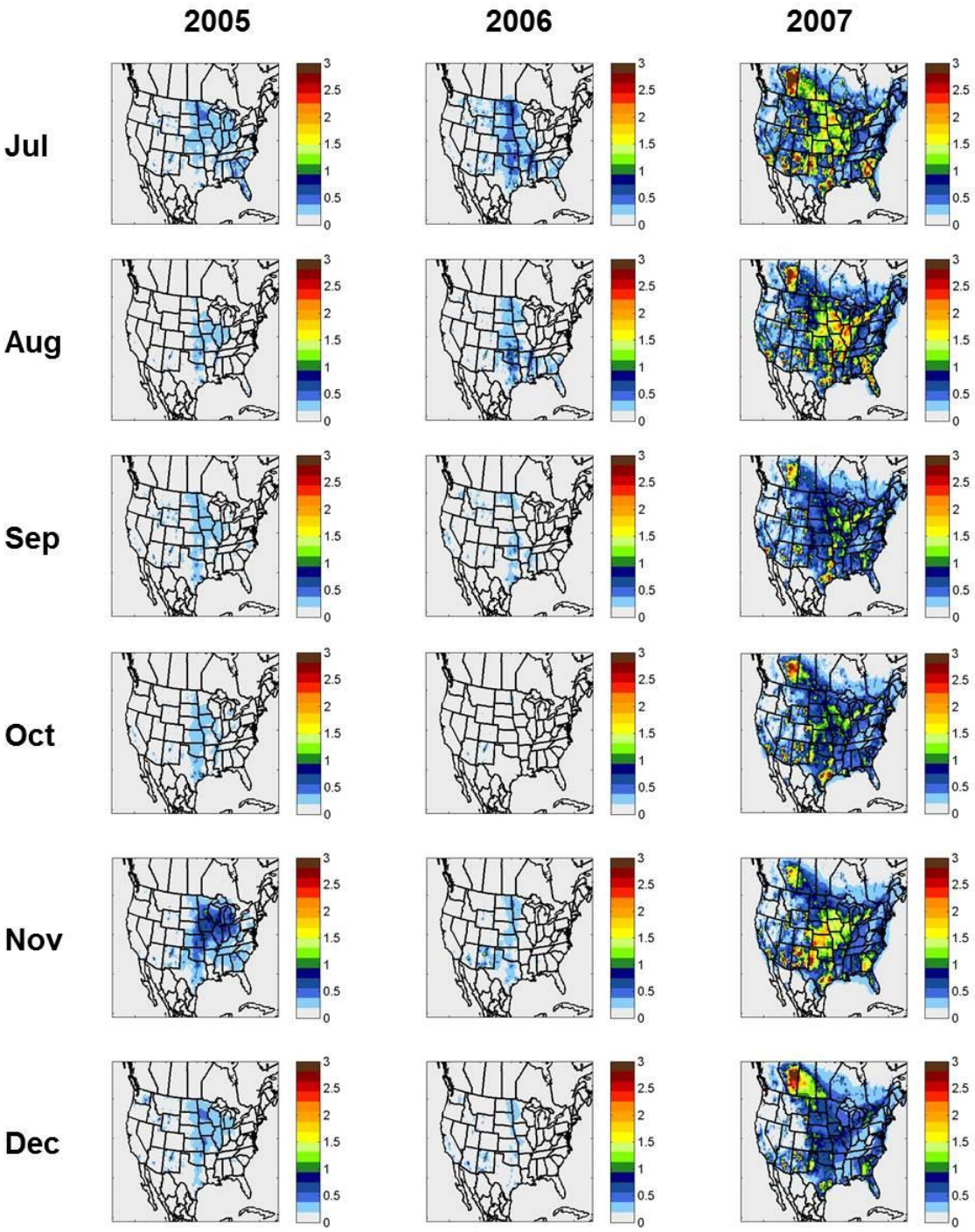

Figure B.4. Continued

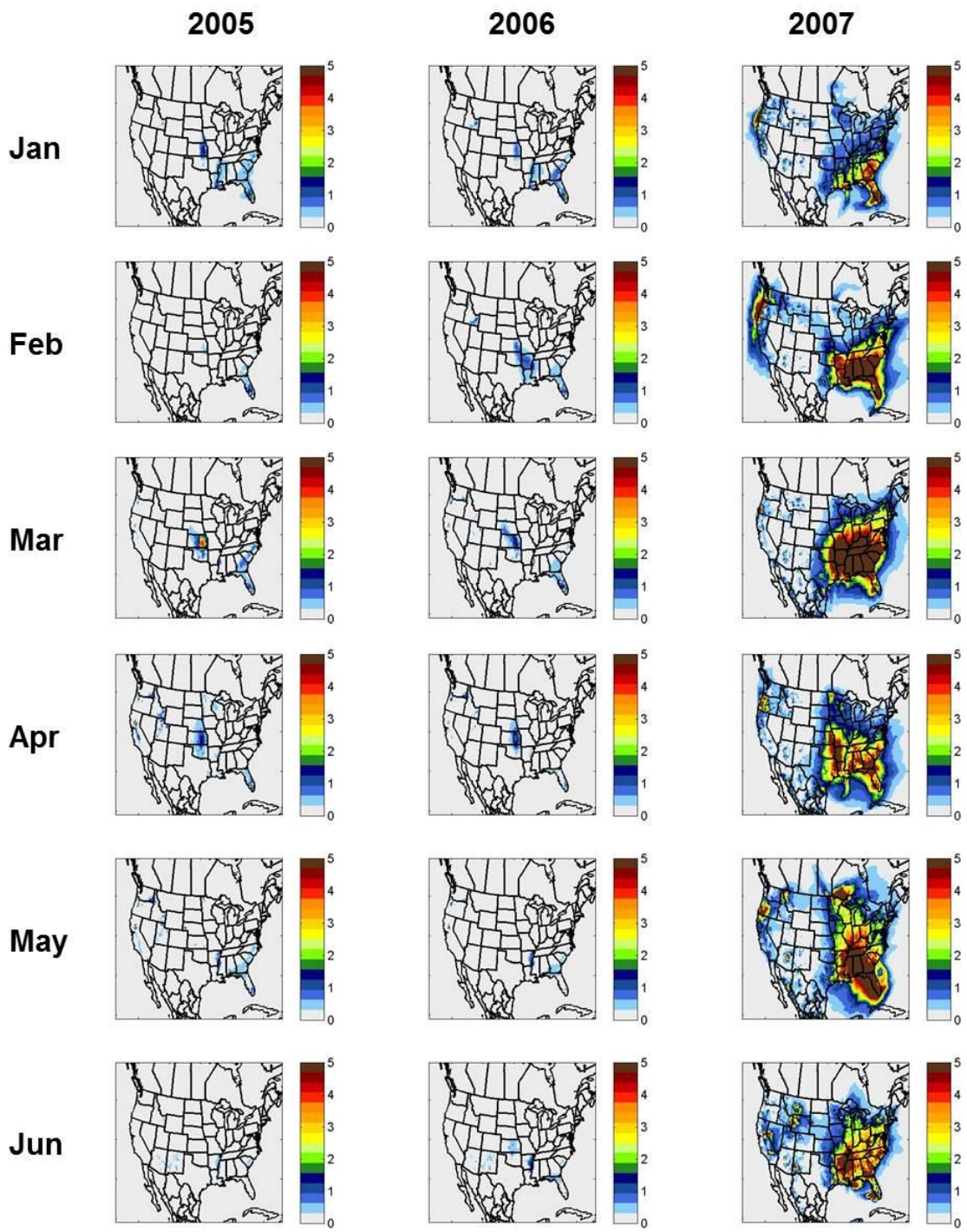

Figure B.5. Spatial fields of monthly-averaged, optimized open fire impacts on $PM_{2.5}$ for years 2005-2007.

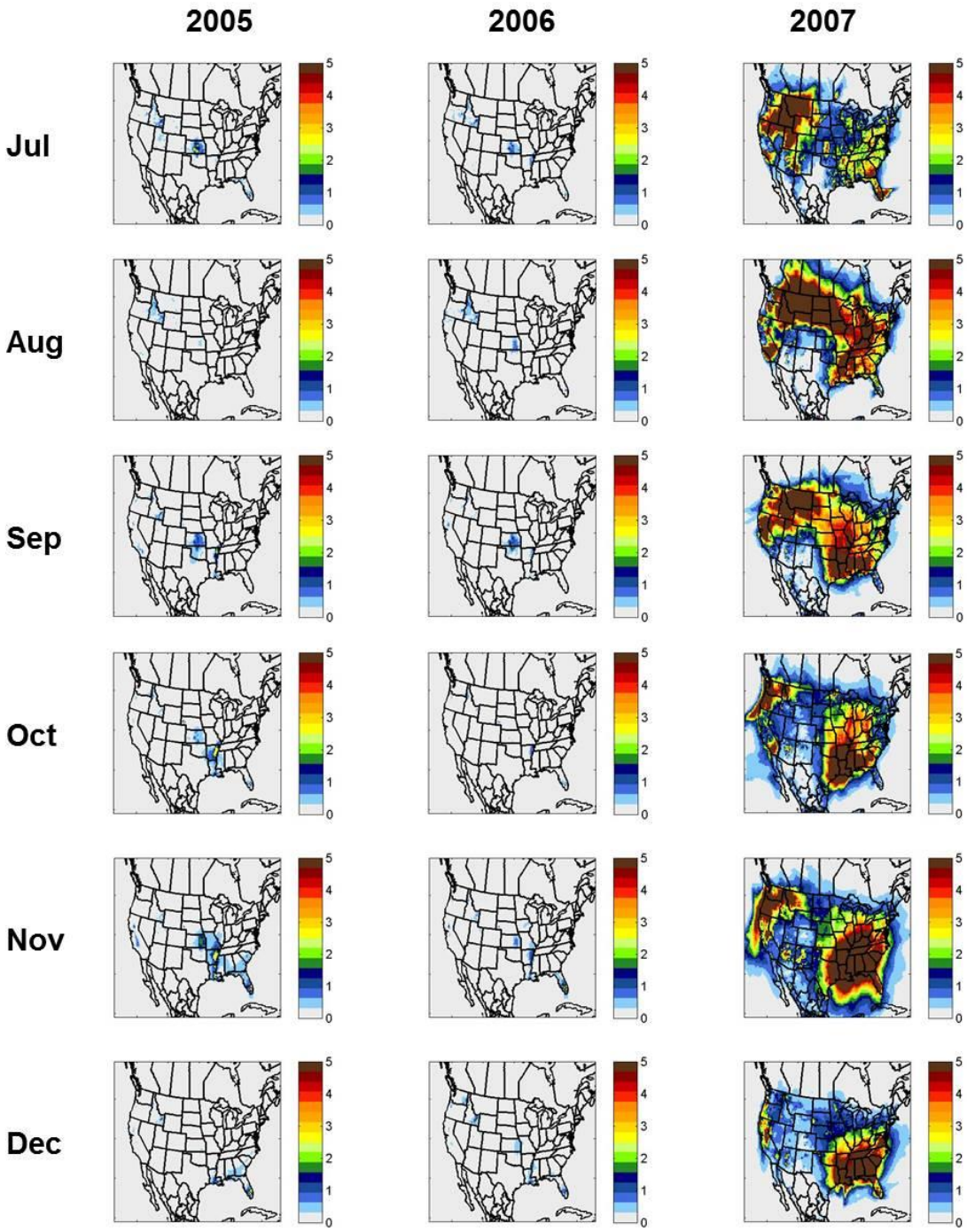

Figure B.5. Continued

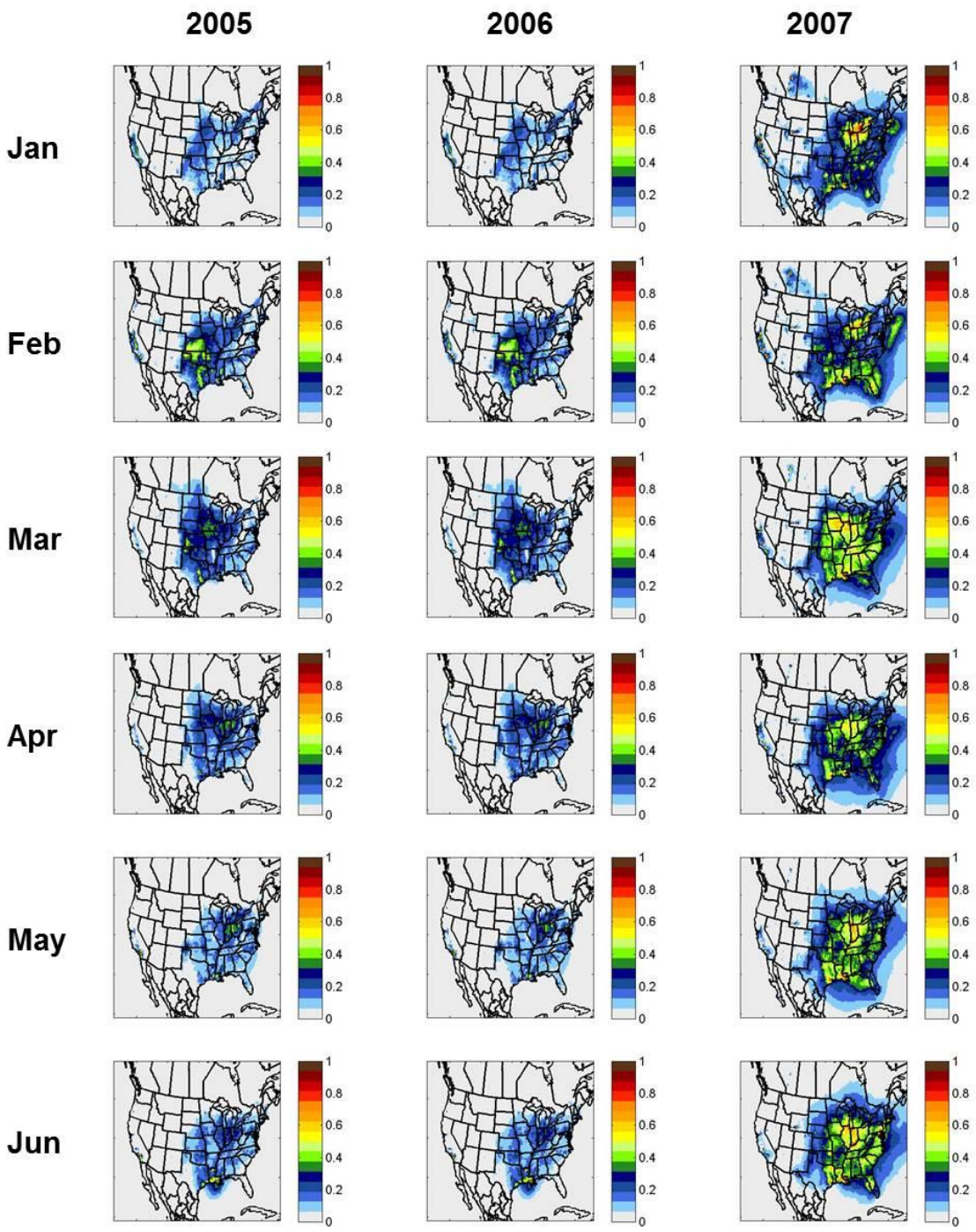

Figure B.6. Spatial fields of monthly-averaged, optimized natural gas impacts on $PM_{2.5}$ for years 2005-2007.

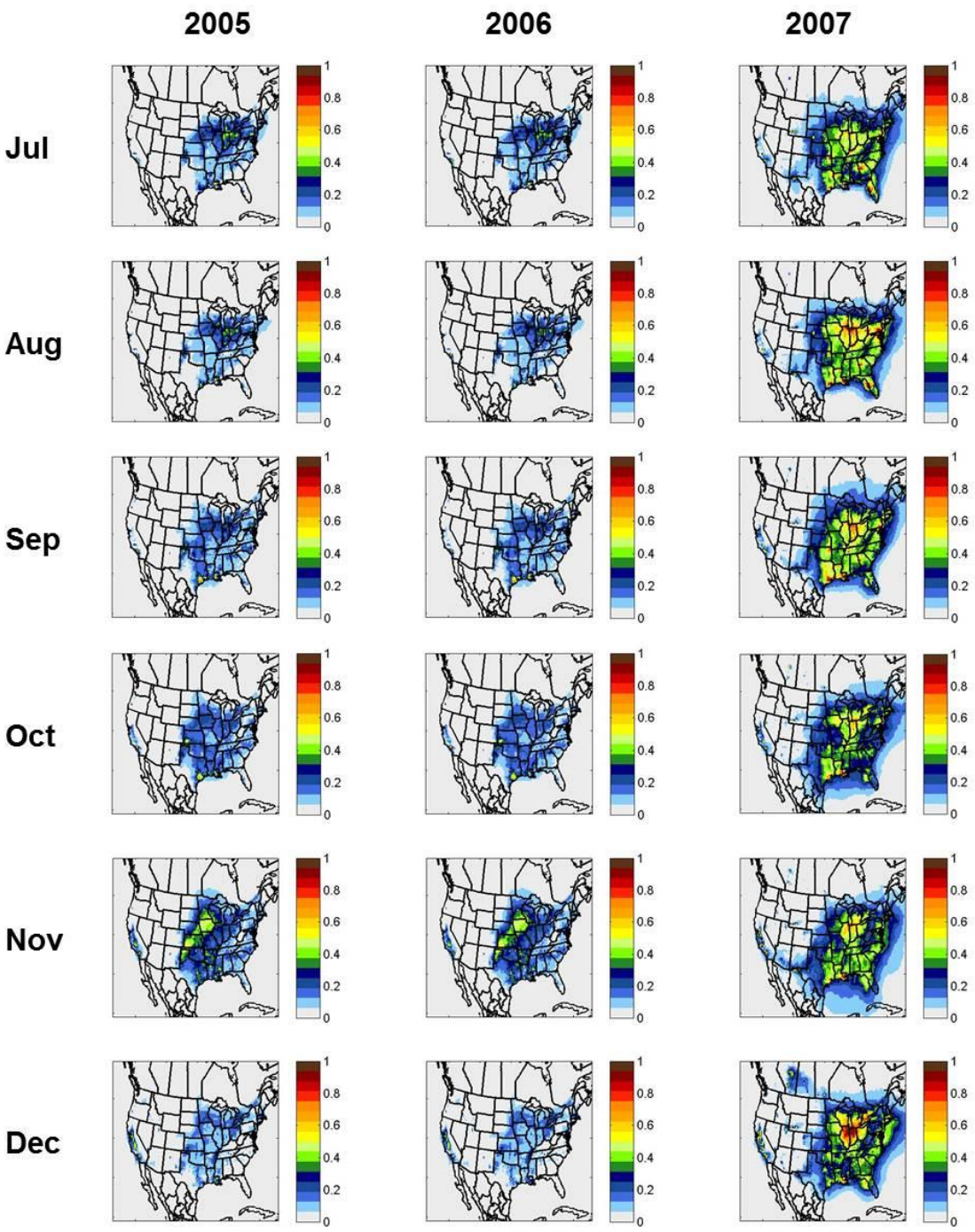

Figure B.6. Continued

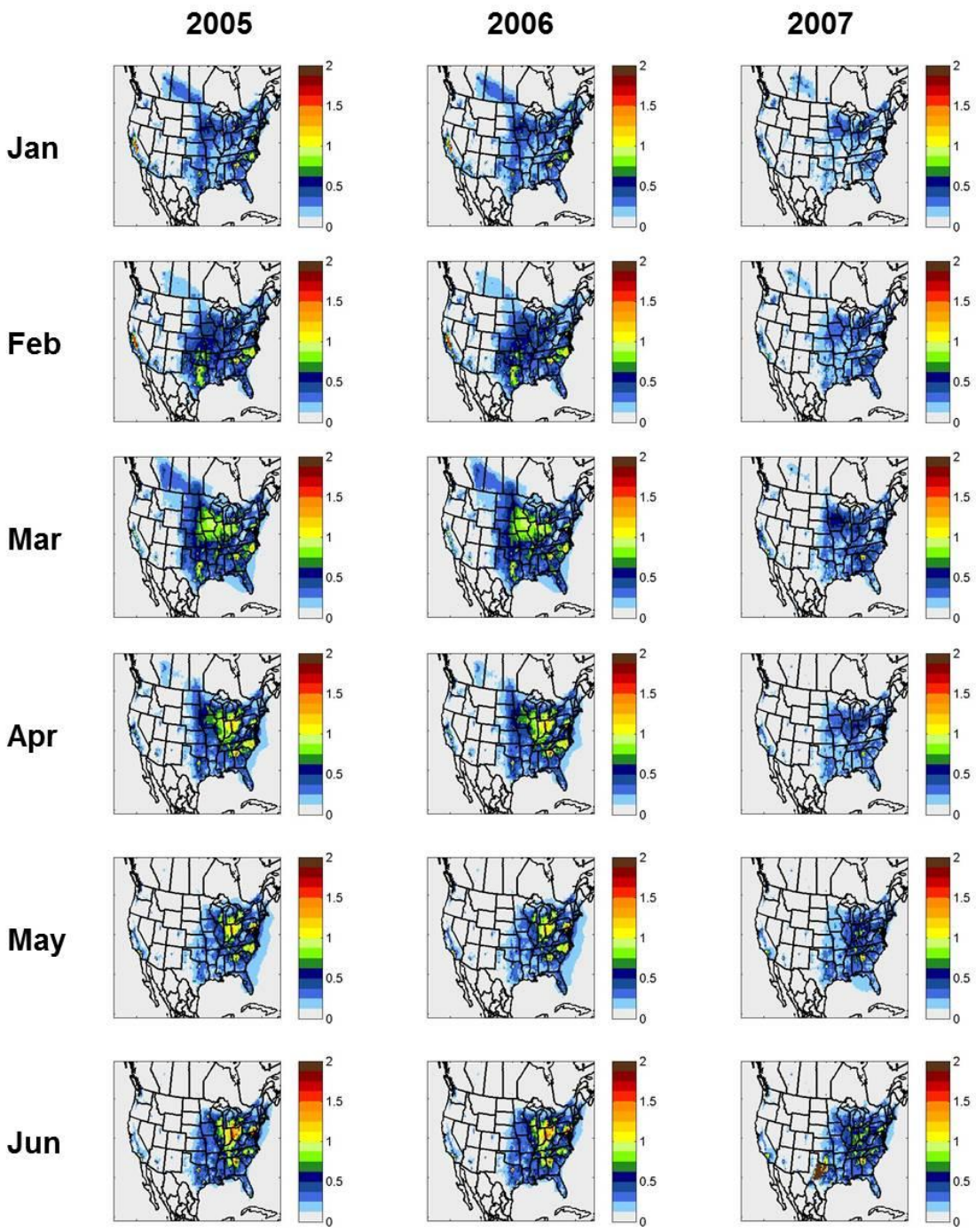

Figure B.7. Spatial fields of monthly-averaged, optimized on-road diesel vehicle impacts on $PM_{2.5}$ for years 2005-2007.

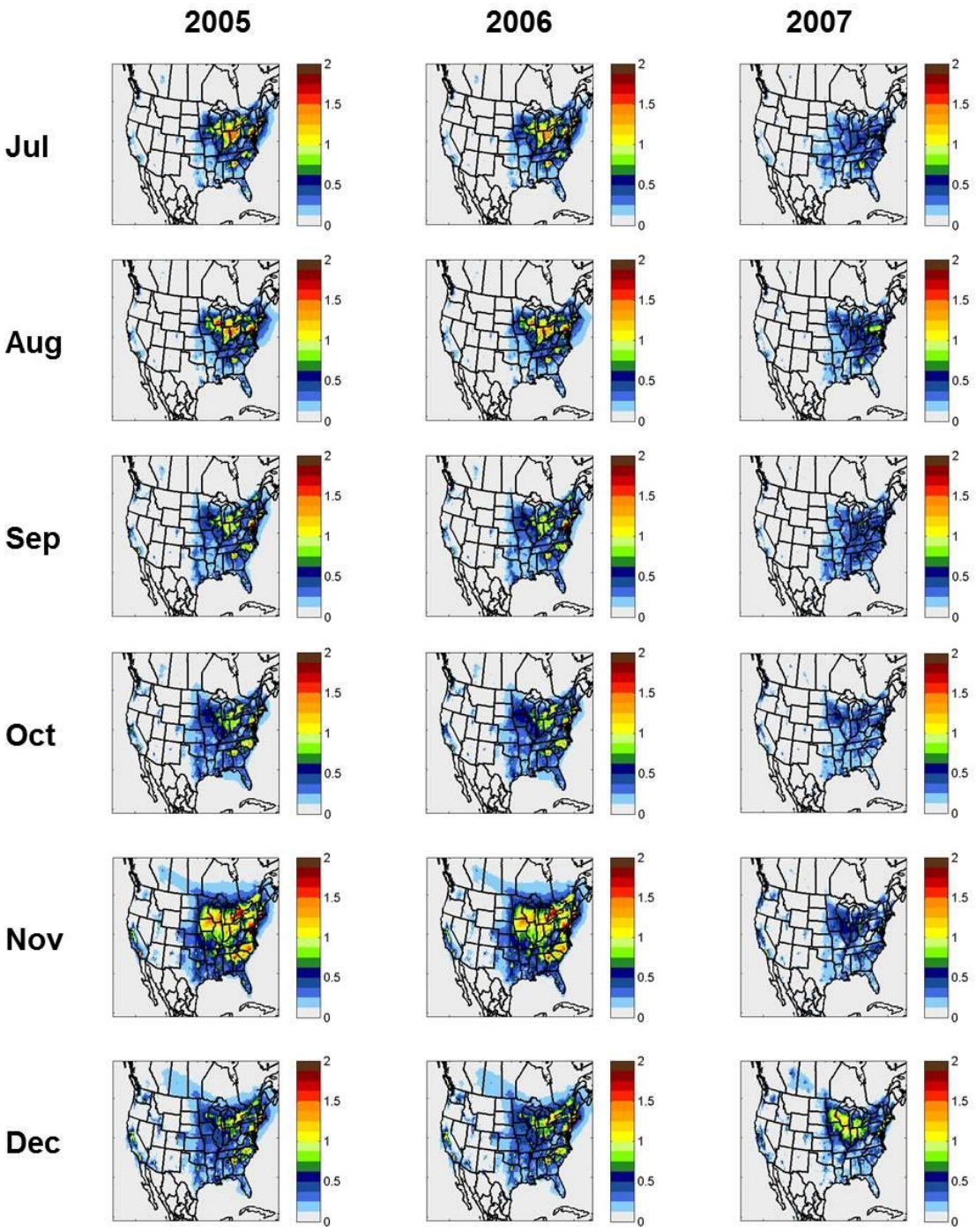

Figure B.7. Continued

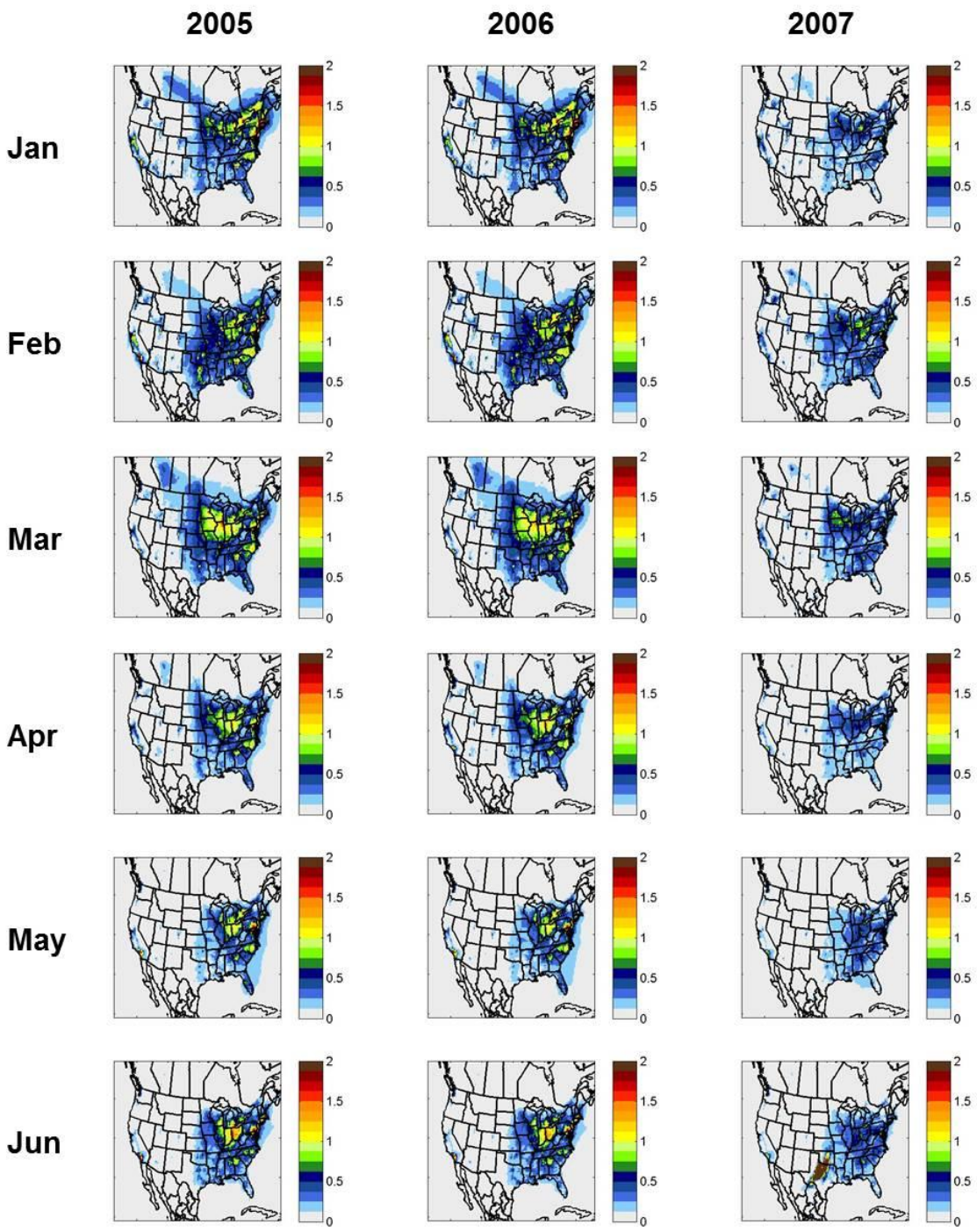

Figure B.8. Spatial fields of monthly-averaged, optimized on-road gasoline vehicle impacts on $PM_{2.5}$ for years 2005-2007.

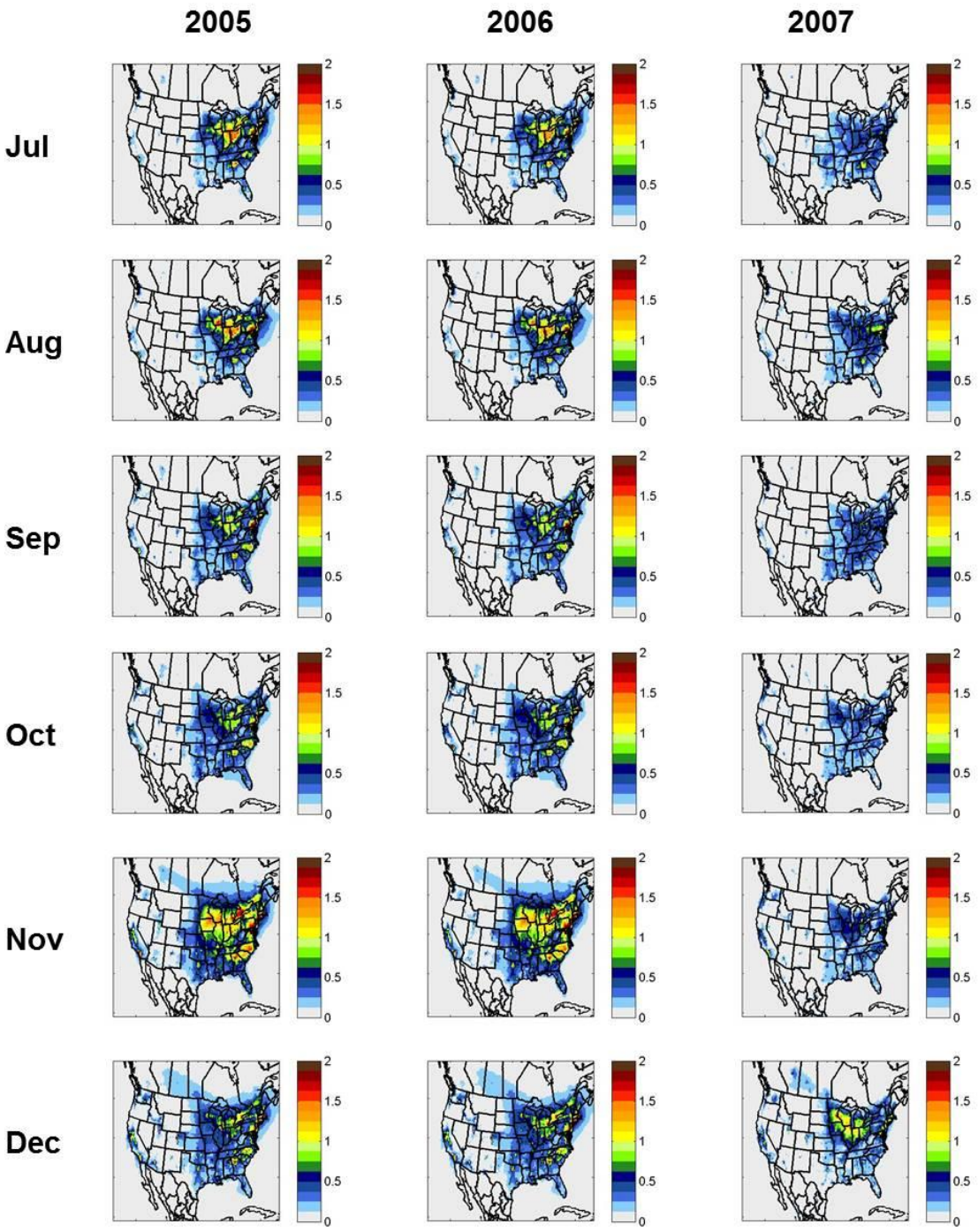

Figure B.8. Continued.

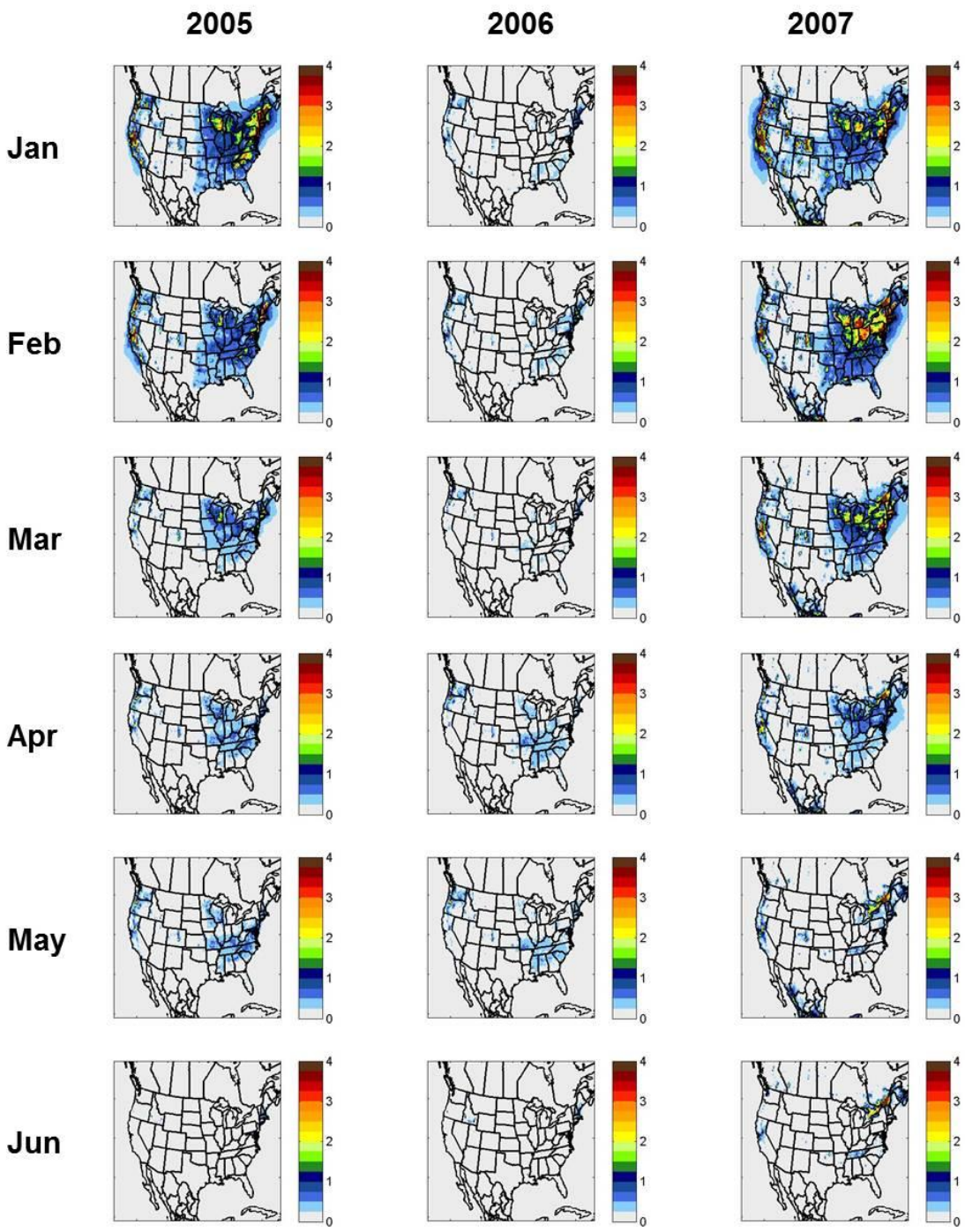

Figure B.9. Spatial fields of monthly-averaged, optimized wood burning impacts on $PM_{2.5}$ for years 2005-2007.

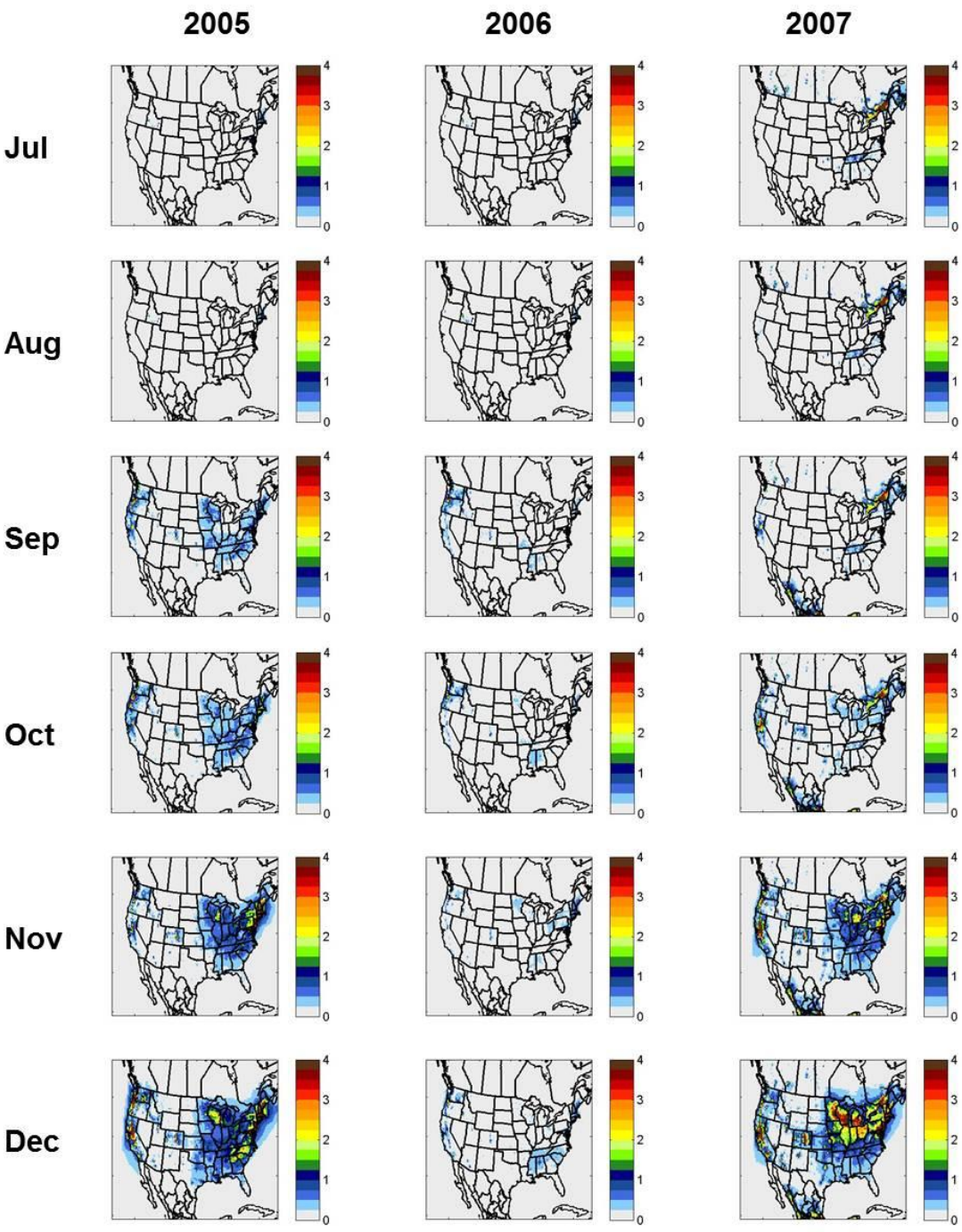

Figure B.9. Continued.

APPENDIX C

CHAPTER 4 SUPPLEMENTAL INFORMATION

Table C.1. Number of CSN monitors with complete speciated $PM_{2.5}$ data in each EPA region during the 2006 study.

region number	number of monitors
1	8
2	5
3	28
4	56
5	42
6	11
7	4
8	13
9	14
10	7
total	**188**

Table C.2. $PM_{2.5}$ source profiles derived from Reff et al. (2009). Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	1.8E-01	0.0E+00	3.9E-02	4.9E-02	4.8E-01	1.1E-01	6.5E-01	7.1E-02	4.8E-01
NO3	0.0E+00	7.7E-01	0.0E+00	2.1E-02	1.7E-03	1.0E-01	4.9E-02	3.4E-02	1.4E-02	8.5E-02
NH4	0.0E+00	1.1E-03	0.0E+00	2.0E-03	1.8E-03	5.6E-03	0.0E+00	4.6E-03	5.2E-03	3.3E-02
SO4	0.0E+00	0.0E+00	0.0E+00	2.4E-02	8.6E-04	6.2E-03	0.0E+00	0.0E+00	4.1E-04	4.2E-03
EC	0.0E+00	3.0E-03	0.0E+00	1.2E-01	6.0E-03	8.4E-03	3.3E-01	2.6E-03	1.8E-01	1.2E-01
Na	0.0E+00	0.0E+00	0.0E+00	1.2E-03	1.1E-03	4.1E-03	0.0E+00	3.2E-03	3.1E-02	2.0E-02
Al	0.0E+00	0.0E+00	0.0E+00	4.4E-02	5.9E-02	5.5E-04	0.0E+00	4.4E-04	3.5E-02	2.1E-03
Si	0.0E+00	0.0E+00	0.0E+00	8.0E-02	1.7E-01	1.0E-03	3.5E-03	8.1E-04	4.0E-02	2.6E-03
Cl	0.0E+00	2.1E-04	0.0E+00	6.0E-03	1.4E-03	2.3E-02	1.8E-03	8.8E-03	1.0E-01	3.7E-02
K	0.0E+00	3.8E-05	0.0E+00	4.3E-03	1.7E-02	1.7E-02	1.7E-04	2.7E-03	3.7E-02	1.6E-03
Ca	0.0E+00	5.8E-04	0.0E+00	3.9E-02	5.4E-02	2.3E-03	2.9E-04	5.2E-04	1.3E-02	9.8E-03
Ti	0.0E+00	4.0E-06	0.0E+00	2.4E-03	3.9E-03	3.3E-04	3.6E-04	2.8E-05	8.9E-04	9.4E-04
Mn	0.0E+00	0.0E+00	0.0E+00	1.8E-04	1.2E-03	6.3E-05	2.1E-05	6.7E-05	5.1E-03	1.9E-04
Fe	0.0E+00	2.6E-04	0.0E+00	2.1E-02	4.4E-02	4.4E-04	5.6E-04	7.7E-04	7.5E-02	8.5E-04
Cu	0.0E+00	0.0E+00	0.0E+00	1.5E-04	1.1E-04	1.9E-05	8.7E-05	1.2E-04	5.5E-03	0.0E+00
Zn	0.0E+00	6.8E-04	0.0E+00	5.1E-04	5.2E-04	1.1E-03	7.1E-04	2.6E-04	2.6E-03	0.0E+00
As	0.0E+00	3.0E-06	0.0E+00	1.2E-06	1.8E-05	6.2E-05	0.0E+00	2.9E-06	4.1E-04	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	2.8E-03	1.1E-06	3.0E-06	2.8E-05	1.0E-05	1.7E-04	5.8E-04
Br	0.0E+00	1.5E-05	0.0E+00	2.9E-04	1.2E-05	2.6E-04	0.0E+00	1.7E-04	1.4E-03	3.8E-04
Sn	0.0E+00	2.5E-05	0.0E+00	1.3E-04	2.8E-05	9.8E-05	0.0E+00	1.2E-05	3.8E-04	0.0E+00
Sb	0.0E+00	1.9E-05	0.0E+00	8.1E-06	3.2E-05	3.5E-06	1.3E-04	2.3E-05	8.2E-04	0.0E+00
Ba	0.0E+00	3.5E-04	0.0E+00	3.9E-03	7.9E-04	1.4E-05	0.0E+00	9.2E-04	2.8E-03	0.0E+00
Pb	0.0E+00	1.0E-06	0.0E+00	1.2E-04	2.2E-04	7.7E-05	0.0E+00	3.8E-04	2.1E-03	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	1.8E-01	4.8E-01	2.1E-01	1.8E-01	4.1E-01	1.1E-01	2.4E-01	2.7E-01	0.0E+00	4.2E-01
NO3	7.7E-01	1.2E-01	9.3E-02	7.5E-01	1.5E-01	2.6E-02	1.0E-01	4.7E-03	0.0E+00	4.5E-02
NH4	1.1E-03	7.0E-04	1.3E-02	1.2E-03	1.5E-03	9.2E-03	8.8E-03	3.6E-04	0.0E+00	7.8E-04
SO4	0.0E+00	0.0E+00	1.7E-03	1.7E-04	1.0E-02	7.2E-03	3.2E-02	7.5E-04	0.0E+00	6.1E-04
EC	3.0E-03	5.0E-04	2.9E-01	3.9E-03	1.8E-02	1.4E-01	1.3E-01	2.8E-02	7.6E-02	4.0E-02
Na	0.0E+00	0.0E+00	8.1E-03	2.0E-05	7.6E-04	3.8E-02	2.2E-03	2.4E-05	3.0E-01	1.3E-03
Al	0.0E+00	0.0E+00	8.3E-04	3.1E-05	1.3E-03	1.3E-02	8.1E-03	2.5E-02	0.0E+00	4.5E-05
Si	0.0E+00	1.2E-03	1.1E-03	1.3E-03	2.9E-02	3.8E-02	6.7E-02	2.5E-02	0.0E+00	8.0E-02
Cl	2.1E-04	0.0E+00	1.6E-02	3.4E-04	1.8E-03	2.1E-02	2.8E-02	3.2E-03	5.4E-01	4.4E-03
K	3.8E-05	0.0E+00	6.5E-04	4.5E-05	2.1E-04	1.7E-02	3.6E-02	2.0E-03	1.1E-02	5.6E-02
Ca	5.8E-04	3.0E-04	4.2E-03	7.1E-04	4.8E-03	2.6E-02	1.0E-02	2.5E-02	1.2E-02	3.8E-03
Ti	4.0E-06	0.0E+00	7.3E-04	5.8E-05	1.1E-03	4.3E-03	6.1E-04	1.7E-01	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	9.6E-05	1.6E-05	3.5E-04	4.9E-04	1.7E-04	0.0E+00	0.0E+00	0.0E+00
Fe	2.6E-04	1.0E-04	5.2E-04	1.9E-03	3.7E-02	8.0E-03	8.6E-03	3.3E-04	0.0E+00	3.7E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.7E-04	3.6E-03	2.3E-04	2.3E-03	1.5E-05	0.0E+00	0.0E+00
Zn	6.8E-04	3.0E-04	9.5E-05	7.6E-04	2.7E-03	6.8E-04	7.5E-03	1.1E-05	0.0E+00	9.0E-05
As	3.0E-06	0.0E+00	2.4E-07	3.1E-06	6.0E-06	3.5E-05	2.3E-05	0.0E+00	0.0E+00	8.2E-07
Se	0.0E+00	0.0E+00	2.6E-04	5.8E-07	1.2E-05	3.0E-05	1.9E-05	0.0E+00	0.0E+00	0.0E+00
Br	1.5E-05	0.0E+00	1.5E-04	1.5E-05	4.9E-05	4.3E-04	5.6E-05	9.9E-05	0.0E+00	8.2E-06
Sn	2.5E-05	0.0E+00	2.0E-06	1.1E-04	2.0E-03	7.9E-05	2.1E-05	0.0E+00	0.0E+00	4.1E-06
Sb	1.9E-05	2.0E-04	1.3E-04	2.0E-05	1.1E-04	1.5E-04	4.1E-05	4.7E-03	0.0E+00	0.0E+00
Ba	3.5E-04	0.0E+00	2.8E-05	8.1E-04	1.1E-02	6.2E-04	5.1E-04	1.3E-02	0.0E+00	0.0E+00
Pb	1.0E-06	3.0E-04	3.3E-07	7.6E-06	3.2E-04	2.2E-04	2.3E-05	3.0E-04	0.0E+00	2.5E-06

Table C.3. Winter (December, January, February) $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	1.8E-01	0.0E+00	3.9E-02	4.9E-02	4.8E-01	1.1E-01	6.5E-01	7.0E-02	4.8E-01
NO3	0.0E+00	1.1E-03	0.0E+00	2.0E-03	1.8E-03	5.6E-03	0.0E+00	4.6E-03	5.2E-03	3.3E-02
NH4	0.0E+00	0.0E+00	0.0E+00	2.4E-02	8.6E-04	6.2E-03	0.0E+00	0.0E+00	4.1E-04	4.2E-03
SO4	0.0E+00	3.0E-03	0.0E+00	1.2E-01	6.0E-03	8.4E-03	3.3E-01	2.6E-03	1.8E-01	1.2E-01
EC	0.0E+00	7.1E-01	0.0E+00	2.0E-02	1.7E-03	9.5E-02	4.7E-02	3.3E-02	1.3E-02	8.1E-02
Na	0.0E+00	0.0E+00	0.0E+00	1.1E-03	1.1E-03	3.9E-03	0.0E+00	3.1E-03	3.0E-02	2.0E-02
Al	0.0E+00	0.0E+00	0.0E+00	4.5E-02	6.0E-02	5.3E-04	0.0E+00	4.2E-04	3.5E-02	2.0E-03
Si	0.0E+00	0.0E+00	0.0E+00	9.6E-02	1.9E-01	9.8E-04	3.4E-03	7.8E-04	4.7E-02	2.6E-03
Cl	0.0E+00	2.0E-04	0.0E+00	5.6E-03	1.3E-03	2.2E-02	1.7E-03	8.5E-03	6.2E-02	3.4E-02
K	0.0E+00	3.6E-05	0.0E+00	5.2E-03	2.0E-02	1.7E-02	1.7E-04	2.8E-03	5.3E-02	1.6E-03
Ca	0.0E+00	5.7E-04	0.0E+00	5.4E-02	6.4E-02	2.3E-03	2.8E-04	5.1E-04	1.6E-02	1.1E-02
Ti	0.0E+00	3.8E-06	0.0E+00	2.5E-03	3.8E-03	3.2E-04	3.5E-04	2.7E-05	8.6E-04	9.2E-04
Mn	0.0E+00	0.0E+00	0.0E+00	1.7E-04	1.2E-03	6.1E-05	2.0E-05	6.4E-05	5.9E-03	1.8E-04
Fe	0.0E+00	2.5E-04	0.0E+00	3.1E-02	5.5E-02	4.2E-04	5.5E-04	7.8E-04	1.1E-01	8.4E-04
Cu	0.0E+00	0.0E+00	0.0E+00	1.5E-04	1.1E-04	1.8E-05	8.4E-05	1.2E-04	6.7E-03	0.0E+00
Zn	0.0E+00	7.0E-04	0.0E+00	5.6E-04	5.2E-04	1.1E-03	7.4E-04	2.6E-04	3.1E-03	0.0E+00
As	0.0E+00	2.9E-06	0.0E+00	1.2E-06	1.7E-05	6.0E-05	0.0E+00	2.8E-06	4.0E-04	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	3.0E-03	1.1E-06	2.9E-06	2.7E-05	9.7E-06	1.6E-04	5.6E-04
Br	0.0E+00	1.4E-05	0.0E+00	3.1E-04	1.2E-05	2.6E-04	0.0E+00	1.7E-04	1.6E-03	3.8E-04
Sn	0.0E+00	2.4E-05	0.0E+00	1.2E-04	2.7E-05	9.4E-05	0.0E+00	1.1E-05	3.6E-04	0.0E+00
Sb	0.0E+00	1.8E-05	0.0E+00	7.8E-06	3.1E-05	3.4E-06	1.3E-04	2.2E-05	7.8E-04	0.0E+00
Ba	0.0E+00	3.4E-04	0.0E+00	3.8E-03	7.6E-04	1.3E-05	0.0E+00	8.8E-04	2.7E-03	0.0E+00
Pb	0.0E+00	9.6E-07	0.0E+00	1.2E-04	2.1E-04	7.3E-05	0.0E+00	3.7E-04	2.2E-03	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	1.8E-01	4.8E-01	2.1E-01	1.8E-01	4.1E-01	1.1E-01	2.4E-01	2.7E-01	0.0E+00	4.2E-01
NO3	1.1E-03	7.0E-04	1.3E-02	1.2E-03	1.5E-03	9.2E-03	8.8E-03	3.5E-04	0.0E+00	7.8E-04
NH4	0.0E+00	0.0E+00	1.7E-03	1.7E-04	1.0E-02	7.2E-03	3.2E-02	7.5E-04	0.0E+00	6.1E-04
SO4	3.0E-03	5.0E-04	2.9E-01	3.9E-03	1.8E-02	1.4E-01	1.3E-01	2.8E-02	7.6E-02	4.0E-02
EC	7.6E-01	1.2E-01	8.8E-02	7.3E-01	1.4E-01	2.5E-02	9.6E-02	4.5E-03	0.0E+00	4.3E-02
Na	0.0E+00	0.0E+00	7.7E-03	1.9E-05	7.3E-04	4.0E-02	2.1E-03	2.3E-05	3.0E-01	1.2E-03
Al	0.0E+00	0.0E+00	8.0E-04	3.0E-05	1.3E-03	1.5E-02	7.8E-03	2.4E-02	0.0E+00	4.3E-05
Si	0.0E+00	1.2E-03	1.0E-03	1.4E-03	3.7E-02	5.0E-02	7.2E-02	2.6E-02	0.0E+00	1.0E-01
Cl	2.0E-04	0.0E+00	1.6E-02	3.3E-04	1.7E-03	1.6E-02	2.6E-02	3.0E-03	3.1E-01	4.2E-03
K	3.7E-05	0.0E+00	6.2E-04	4.4E-05	2.0E-04	2.6E-02	4.2E-02	1.9E-03	1.2E-02	8.8E-02
Ca	6.5E-04	2.9E-04	4.0E-03	8.5E-04	6.1E-03	3.8E-02	1.0E-02	2.9E-02	1.2E-02	4.8E-03
Ti	3.8E-06	0.0E+00	7.0E-04	5.7E-05	1.1E-03	4.7E-03	5.8E-04	1.8E-01	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	9.2E-05	1.6E-05	3.6E-04	5.4E-04	1.7E-04	0.0E+00	0.0E+00	0.0E+00
Fe	2.8E-04	9.6E-05	5.0E-04	2.8E-03	5.9E-02	1.2E-02	9.1E-03	3.2E-04	0.0E+00	3.6E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.9E-04	4.7E-03	2.4E-04	2.4E-03	1.5E-05	0.0E+00	0.0E+00
Zn	9.3E-04	3.0E-04	9.1E-05	1.1E-03	3.8E-03	8.6E-04	8.6E-03	1.1E-05	0.0E+00	9.1E-05
As	2.9E-06	0.0E+00	2.3E-07	3.0E-06	5.8E-06	3.4E-05	2.2E-05	0.0E+00	0.0E+00	7.9E-07
Se	0.0E+00	0.0E+00	2.5E-04	5.6E-07	1.1E-05	2.9E-05	1.8E-05	0.0E+00	0.0E+00	0.0E+00
Br	1.5E-05	0.0E+00	1.5E-04	1.5E-05	4.8E-05	5.0E-04	5.3E-05	9.6E-05	0.0E+00	7.9E-06
Sn	2.4E-05	0.0E+00	1.9E-06	1.1E-04	2.0E-03	7.6E-05	2.0E-05	0.0E+00	0.0E+00	3.9E-06
Sb	1.8E-05	1.9E-04	1.2E-04	1.9E-05	1.1E-04	1.4E-04	3.9E-05	4.5E-03	0.0E+00	0.0E+00
Ba	3.4E-04	0.0E+00	2.7E-05	8.0E-04	1.1E-02	6.0E-04	4.9E-04	1.2E-02	0.0E+00	0.0E+00
Pb	9.6E-07	2.9E-04	3.2E-07	7.3E-06	3.1E-04	2.2E-04	2.2E-05	2.9E-04	0.0E+00	2.4E-06

Table C.4. Winter (December, January, February) standard deviations for $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	7.6E-15	0.0E+00	4.7E-16	2.3E-15	8.6E-15	2.9E-15	3.3E-15	1.4E-16	8.3E-16
NO3	0.0E+00	4.4E-17	0.0E+00	5.0E-17	8.3E-17	1.3E-16	0.0E+00	1.7E-16	9.0E-17	3.0E-16
NH4	0.0E+00	0.0E+00	0.0E+00	4.1E-16	3.2E-17	8.6E-17	0.0E+00	0.0E+00	1.1E-17	1.2E-17
SO4	0.0E+00	1.5E-17	0.0E+00	7.0E-15	2.4E-16	2.9E-16	5.7E-15	4.5E-17	7.4E-15	2.9E-15
EC	0.0E+00	6.1E-02	0.0E+00	7.4E-04	6.0E-05	4.1E-03	1.7E-03	1.2E-03	5.2E-04	3.3E-03
Na	0.0E+00	0.0E+00	0.0E+00	4.6E-05	4.0E-05	1.6E-04	0.0E+00	1.4E-04	3.2E-03	1.8E-03
Al	0.0E+00	0.0E+00	0.0E+00	1.7E-02	1.4E-02	1.9E-05	0.0E+00	1.5E-05	6.6E-03	8.1E-05
Si	0.0E+00	0.0E+00	0.0E+00	5.0E-02	7.9E-02	3.6E-05	1.6E-04	2.6E-05	1.2E-02	1.1E-04
Cl	0.0E+00	7.0E-06	0.0E+00	1.3E-03	7.7E-05	3.1E-03	8.1E-05	1.4E-03	8.3E-02	1.1E-02
K	0.0E+00	1.3E-06	0.0E+00	9.2E-04	4.8E-03	2.4E-03	5.7E-06	2.3E-04	1.2E-02	8.4E-05
Ca	0.0E+00	2.2E-05	0.0E+00	1.9E-02	2.1E-02	1.6E-04	9.1E-06	2.1E-05	3.6E-03	1.8E-03
Ti	0.0E+00	1.4E-07	0.0E+00	4.3E-04	5.2E-04	1.2E-05	1.3E-05	9.6E-07	4.8E-05	4.7E-05
Mn	0.0E+00	0.0E+00	0.0E+00	1.2E-05	1.6E-04	2.2E-06	7.1E-07	2.2E-06	1.2E-03	6.9E-06
Fe	0.0E+00	8.5E-06	0.0E+00	7.0E-03	1.6E-02	1.8E-05	1.9E-05	3.5E-05	2.7E-02	3.7E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.1E-05	5.7E-06	6.4E-07	3.1E-06	5.5E-06	1.6E-03	0.0E+00
Zn	0.0E+00	6.0E-05	0.0E+00	7.5E-05	5.6E-05	8.9E-05	5.8E-05	1.3E-05	5.3E-04	0.0E+00
As	0.0E+00	1.0E-07	0.0E+00	4.3E-08	6.1E-07	2.1E-06	0.0E+00	1.0E-07	2.8E-05	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	6.1E-04	3.8E-08	1.0E-07	9.7E-07	3.5E-07	6.2E-06	2.5E-05
Br	0.0E+00	5.1E-07	0.0E+00	4.0E-05	4.3E-07	1.4E-05	0.0E+00	9.3E-06	2.7E-04	2.6E-05
Sn	0.0E+00	8.6E-07	0.0E+00	4.5E-06	9.7E-07	3.4E-06	0.0E+00	4.0E-07	1.3E-05	0.0E+00
Sb	0.0E+00	6.6E-07	0.0E+00	2.8E-07	1.1E-06	1.2E-07	4.6E-06	7.8E-07	2.8E-05	0.0E+00
Ba	0.0E+00	1.2E-05	0.0E+00	2.6E-04	2.8E-05	4.8E-07	0.0E+00	3.4E-05	1.1E-04	0.0E+00
Pb	0.0E+00	3.5E-08	0.0E+00	5.1E-06	9.2E-06	2.6E-06	0.0E+00	1.7E-05	2.3E-04	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	7.6E-15	1.6E-14	4.7E-16	7.4E-15	2.2E-14	6.4E-15	3.7E-15	8.3E-16	0.0E+00	2.0E-14
NO3	4.4E-17	1.6E-17	1.9E-16	1.1E-17	2.1E-17	1.7E-16	3.3E-16	5.2E-18	0.0E+00	2.6E-17
NH4	0.0E+00	0.0E+00	5.0E-17	4.9E-19	2.3E-16	3.2E-16	6.1E-16	2.0E-17	0.0E+00	1.5E-17
SO4	1.5E-17	2.9E-17	7.8E-16	3.9E-17	5.1E-16	8.9E-16	2.4E-15	5.8E-16	1.0E-15	2.2E-15
EC	1.2E-01	4.6E-03	3.5E-03	1.9E-01	9.5E-03	9.5E-04	3.9E-03	1.6E-04	0.0E+00	2.0E-03
Na	0.0E+00	0.0E+00	2.8E-04	6.9E-07	2.8E-05	1.6E-02	7.7E-05	8.3E-07	1.5E-01	5.3E-05
Al	0.0E+00	0.0E+00	2.9E-05	1.1E-06	6.6E-05	3.0E-03	4.2E-04	1.7E-03	0.0E+00	1.5E-06
Si	0.0E+00	3.8E-05	3.6E-05	2.2E-04	1.2E-02	1.8E-02	1.3E-02	2.5E-03	0.0E+00	5.0E-02
Cl	1.1E-05	0.0E+00	6.3E-04	2.9E-05	1.7E-04	1.2E-02	5.0E-03	1.2E-04	3.7E-01	8.9E-04
K	1.2E-06	0.0E+00	2.2E-05	1.4E-06	6.9E-06	5.2E-03	7.6E-03	6.6E-05	4.7E-03	2.0E-02
Ca	8.5E-05	9.3E-06	1.4E-04	1.5E-04	1.3E-03	1.1E-02	1.3E-03	5.0E-03	6.4E-03	9.6E-04
Ti	1.4E-07	0.0E+00	2.5E-05	2.9E-06	1.5E-04	1.2E-03	2.2E-05	6.8E-02	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	3.3E-06	6.5E-07	3.7E-05	7.7E-05	6.0E-06	0.0E+00	0.0E+00	0.0E+00
Fe	2.0E-05	3.2E-06	1.8E-05	5.4E-04	1.3E-02	2.6E-03	1.0E-03	1.1E-05	0.0E+00	1.1E-06
Cu	0.0E+00	0.0E+00	0.0E+00	3.3E-05	1.1E-03	2.7E-05	2.8E-04	5.3E-07	0.0E+00	0.0E+00
Zn	1.8E-04	1.1E-05	3.3E-06	2.2E-04	7.4E-04	1.3E-04	1.5E-03	3.8E-07	0.0E+00	4.6E-06
As	1.0E-07	0.0E+00	8.3E-09	1.0E-07	2.0E-07	1.4E-06	7.8E-07	0.0E+00	0.0E+00	2.8E-08
Se	0.0E+00	0.0E+00	8.9E-06	2.0E-08	4.0E-07	1.0E-06	6.5E-07	0.0E+00	0.0E+00	0.0E+00
Br	6.5E-07	0.0E+00	5.2E-06	7.8E-07	2.2E-06	8.3E-05	1.9E-06	3.3E-06	0.0E+00	2.6E-07
Sn	8.6E-07	0.0E+00	6.9E-08	3.8E-06	7.4E-05	2.7E-06	7.3E-07	0.0E+00	0.0E+00	1.4E-07
Sb	6.5E-07	6.9E-06	4.4E-06	6.7E-07	3.9E-06	5.0E-06	1.4E-06	1.6E-04	0.0E+00	0.0E+00
Ba	1.6E-05	0.0E+00	9.7E-07	5.5E-05	1.2E-03	2.4E-05	1.8E-05	5.4E-04	0.0E+00	0.0E+00
Pb	3.4E-08	1.0E-05	1.1E-08	2.5E-07	1.9E-05	1.2E-05	7.8E-07	1.0E-05	0.0E+00	8.4E-08

Table C.5. Spring (March, April, May) $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	1.8E-01	0.0E+00	3.9E-02	4.9E-02	4.8E-01	1.1E-01	6.5E-01	7.0E-02	4.8E-01
NO3	0.0E+00	1.1E-03	0.0E+00	2.0E-03	1.8E-03	5.6E-03	0.0E+00	4.6E-03	5.2E-03	3.3E-02
NH4	0.0E+00	0.0E+00	0.0E+00	2.4E-02	8.6E-04	6.2E-03	0.0E+00	0.0E+00	4.1E-04	4.2E-03
SO4	0.0E+00	3.0E-03	0.0E+00	1.2E-01	6.0E-03	8.4E-03	3.3E-01	2.6E-03	1.8E-01	1.2E-01
EC	0.0E+00	6.9E-01	0.0E+00	1.9E-02	1.6E-03	9.5E-02	4.7E-02	3.2E-02	1.3E-02	8.1E-02
Na	0.0E+00	0.0E+00	0.0E+00	1.1E-03	1.1E-03	3.9E-03	0.0E+00	3.1E-03	3.0E-02	1.9E-02
Al	0.0E+00	0.0E+00	0.0E+00	4.6E-02	6.1E-02	5.2E-04	0.0E+00	4.2E-04	3.6E-02	2.0E-03
Si	0.0E+00	0.0E+00	0.0E+00	1.1E-01	2.1E-01	9.7E-04	3.4E-03	7.8E-04	5.1E-02	2.6E-03
Cl	0.0E+00	2.0E-04	0.0E+00	5.7E-03	1.3E-03	2.2E-02	1.7E-03	8.6E-03	7.2E-02	3.5E-02
K	0.0E+00	3.6E-05	0.0E+00	5.3E-03	2.1E-02	1.7E-02	1.7E-04	2.9E-03	5.4E-02	1.6E-03
Ca	0.0E+00	5.7E-04	0.0E+00	5.7E-02	6.7E-02	2.3E-03	2.7E-04	5.1E-04	1.7E-02	1.1E-02
Ti	0.0E+00	3.8E-06	0.0E+00	2.5E-03	3.9E-03	3.2E-04	3.5E-04	2.7E-05	8.6E-04	9.1E-04
Mn	0.0E+00	0.0E+00	0.0E+00	1.7E-04	1.2E-03	6.0E-05	2.0E-05	6.4E-05	5.7E-03	1.8E-04
Fe	0.0E+00	2.5E-04	0.0E+00	3.1E-02	5.7E-02	4.2E-04	5.4E-04	7.7E-04	1.1E-01	8.3E-04
Cu	0.0E+00	0.0E+00	0.0E+00	1.5E-04	1.1E-04	1.8E-05	8.3E-05	1.2E-04	6.7E-03	0.0E+00
Zn	0.0E+00	7.0E-04	0.0E+00	5.7E-04	5.3E-04	1.1E-03	7.3E-04	2.6E-04	3.1E-03	0.0E+00
As	0.0E+00	2.9E-06	0.0E+00	1.2E-06	1.7E-05	5.9E-05	0.0E+00	2.8E-06	4.1E-04	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	3.1E-03	1.0E-06	2.9E-06	2.7E-05	9.6E-06	1.6E-04	5.6E-04
Br	0.0E+00	1.4E-05	0.0E+00	3.0E-04	1.2E-05	2.5E-04	0.0E+00	1.6E-04	1.6E-03	3.7E-04
Sn	0.0E+00	2.4E-05	0.0E+00	1.2E-04	2.7E-05	9.4E-05	0.0E+00	1.1E-05	3.6E-04	0.0E+00
Sb	0.0E+00	1.8E-05	0.0E+00	7.7E-06	3.1E-05	3.4E-06	1.3E-04	2.2E-05	7.8E-04	0.0E+00
Ba	0.0E+00	3.3E-04	0.0E+00	3.8E-03	7.5E-04	1.3E-05	0.0E+00	8.7E-04	2.7E-03	0.0E+00
Pb	0.0E+00	9.5E-07	0.0E+00	1.2E-04	2.1E-04	7.3E-05	0.0E+00	3.6E-04	2.1E-03	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	1.8E-01	4.8E-01	2.1E-01	1.8E-01	4.1E-01	1.1E-01	2.4E-01	2.7E-01	0.0E+00	4.2E-01
NO3	1.1E-03	7.0E-04	1.3E-02	1.2E-03	1.5E-03	9.2E-03	8.8E-03	3.5E-04	0.0E+00	7.8E-04
NH4	0.0E+00	0.0E+00	1.7E-03	1.7E-04	1.0E-02	7.2E-03	3.2E-02	7.5E-04	0.0E+00	6.1E-04
SO4	3.0E-03	5.0E-04	2.9E-01	3.9E-03	1.8E-02	1.4E-01	1.3E-01	2.8E-02	7.6E-02	4.0E-02
EC	7.1E-01	1.2E-01	8.8E-02	6.7E-01	1.4E-01	2.5E-02	9.5E-02	4.4E-03	0.0E+00	4.2E-02
Na	0.0E+00	0.0E+00	7.7E-03	1.9E-05	7.2E-04	4.0E-02	2.1E-03	2.3E-05	3.0E-01	1.2E-03
Al	0.0E+00	0.0E+00	7.9E-04	2.9E-05	1.3E-03	1.5E-02	7.8E-03	2.4E-02	0.0E+00	4.3E-05
Si	0.0E+00	1.2E-03	1.0E-03	1.5E-03	4.0E-02	5.8E-02	7.7E-02	2.6E-02	0.0E+00	1.3E-01
Cl	2.0E-04	0.0E+00	1.6E-02	3.3E-04	1.7E-03	1.8E-02	2.7E-02	3.0E-03	3.4E-01	4.3E-03
K	3.7E-05	0.0E+00	6.2E-04	4.4E-05	2.0E-04	2.6E-02	4.3E-02	1.9E-03	1.2E-02	9.0E-02
Ca	6.7E-04	2.9E-04	4.0E-03	8.7E-04	6.1E-03	4.1E-02	1.1E-02	2.9E-02	1.2E-02	5.0E-03
Ti	3.8E-06	0.0E+00	6.9E-04	5.7E-05	1.1E-03	4.9E-03	5.8E-04	1.9E-01	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	9.1E-05	1.6E-05	3.5E-04	5.2E-04	1.6E-04	0.0E+00	0.0E+00	0.0E+00
Fe	2.8E-04	9.6E-05	5.0E-04	2.7E-03	6.0E-02	1.3E-02	9.1E-03	3.2E-04	0.0E+00	3.5E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.8E-04	4.6E-03	2.4E-04	2.4E-03	1.5E-05	0.0E+00	0.0E+00
Zn	9.6E-04	3.0E-04	9.0E-05	1.1E-03	3.7E-03	8.8E-04	8.7E-03	1.0E-05	0.0E+00	9.0E-05
As	2.9E-06	0.0E+00	2.3E-07	3.0E-06	5.8E-06	3.3E-05	2.2E-05	0.0E+00	0.0E+00	7.8E-07
Se	0.0E+00	0.0E+00	2.5E-04	5.5E-07	1.1E-05	2.9E-05	1.8E-05	0.0E+00	0.0E+00	0.0E+00
Br	1.5E-05	0.0E+00	1.5E-04	1.5E-05	4.7E-05	4.9E-04	5.3E-05	9.5E-05	0.0E+00	7.8E-06
Sn	2.4E-05	0.0E+00	1.9E-06	1.1E-04	1.9E-03	7.5E-05	2.0E-05	0.0E+00	0.0E+00	3.9E-06
Sb	1.8E-05	1.9E-04	1.2E-04	1.9E-05	1.1E-04	1.4E-04	3.9E-05	4.5E-03	0.0E+00	0.0E+00
Ba	3.4E-04	0.0E+00	2.7E-05	7.9E-04	1.1E-02	6.0E-04	4.9E-04	1.2E-02	0.0E+00	0.0E+00
Pb	9.5E-07	2.9E-04	3.2E-07	7.2E-06	3.1E-04	2.2E-04	2.2E-05	2.9E-04	0.0E+00	2.3E-06

Table C.6. Spring (March, April, May) standard deviations for $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	7.6E-15	0.0E+00	4.9E-16	2.3E-15	8.7E-15	2.9E-15	3.3E-15	9.7E-17	8.3E-16
NO3	0.0E+00	4.4E-17	0.0E+00	4.9E-17	8.3E-17	1.3E-16	0.0E+00	1.7E-16	9.3E-17	3.0E-16
NH4	0.0E+00	0.0E+00	0.0E+00	4.1E-16	3.2E-17	8.8E-17	0.0E+00	0.0E+00	1.1E-17	8.7E-18
SO4	0.0E+00	1.6E-17	0.0E+00	7.1E-15	2.4E-16	3.0E-16	5.7E-15	4.6E-17	7.4E-15	2.7E-15
EC	0.0E+00	5.7E-02	0.0E+00	7.3E-04	5.8E-05	3.8E-03	1.7E-03	1.2E-03	5.1E-04	3.2E-03
Na	0.0E+00	0.0E+00	0.0E+00	4.2E-05	3.9E-05	1.5E-04	0.0E+00	1.4E-04	5.2E-03	3.5E-03
Al	0.0E+00	0.0E+00	0.0E+00	1.9E-02	1.7E-02	1.8E-05	0.0E+00	1.4E-05	7.7E-03	7.8E-05
Si	0.0E+00	0.0E+00	0.0E+00	4.1E-02	8.3E-02	3.6E-05	1.7E-04	2.6E-05	1.3E-02	1.3E-04
Cl	0.0E+00	6.9E-06	0.0E+00	1.2E-03	8.7E-05	2.8E-03	8.0E-05	1.3E-03	7.6E-02	8.5E-03
K	0.0E+00	1.2E-06	0.0E+00	1.0E-03	5.7E-03	2.6E-03	5.4E-06	2.6E-04	1.2E-02	9.5E-05
Ca	0.0E+00	2.2E-05	0.0E+00	1.6E-02	2.1E-02	1.5E-04	8.6E-06	1.9E-05	3.7E-03	1.8E-03
Ti	0.0E+00	1.3E-07	0.0E+00	4.6E-04	6.8E-04	1.2E-05	1.2E-05	9.2E-07	5.2E-05	4.6E-05
Mn	0.0E+00	0.0E+00	0.0E+00	1.1E-05	1.6E-04	2.1E-06	6.9E-07	2.2E-06	1.2E-03	6.7E-06
Fe	0.0E+00	8.3E-06	0.0E+00	7.1E-03	1.8E-02	1.8E-05	1.8E-05	3.7E-05	2.6E-02	4.1E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.2E-05	6.2E-06	6.2E-07	2.9E-06	5.1E-06	1.6E-03	0.0E+00
Zn	0.0E+00	6.9E-05	0.0E+00	9.0E-05	7.3E-05	9.7E-05	5.5E-05	1.3E-05	6.0E-04	0.0E+00
As	0.0E+00	1.0E-07	0.0E+00	4.1E-08	6.0E-07	2.1E-06	0.0E+00	9.7E-08	3.6E-05	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	6.8E-04	3.7E-08	1.0E-07	9.4E-07	3.3E-07	6.6E-06	2.7E-05
Br	0.0E+00	4.9E-07	0.0E+00	3.8E-05	4.1E-07	1.4E-05	0.0E+00	8.6E-06	2.5E-04	2.4E-05
Sn	0.0E+00	8.4E-07	0.0E+00	4.3E-06	9.4E-07	3.3E-06	0.0E+00	3.9E-07	1.3E-05	0.0E+00
Sb	0.0E+00	6.4E-07	0.0E+00	2.7E-07	1.1E-06	1.2E-07	4.4E-06	7.6E-07	2.7E-05	0.0E+00
Ba	0.0E+00	1.2E-05	0.0E+00	3.0E-04	2.9E-05	4.7E-07	0.0E+00	3.1E-05	1.1E-04	0.0E+00
Pb	0.0E+00	3.3E-08	0.0E+00	4.5E-06	9.1E-06	2.5E-06	0.0E+00	1.4E-05	2.0E-04	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	7.6E-15	1.6E-14	4.7E-16	7.4E-15	2.2E-14	6.4E-15	3.8E-15	8.3E-16	0.0E+00	2.0E-14
NO3	4.4E-17	1.6E-17	1.9E-16	1.2E-17	2.1E-17	1.7E-16	3.3E-16	5.2E-18	0.0E+00	2.6E-17
NH4	0.0E+00	0.0E+00	5.1E-17	4.9E-19	2.4E-16	3.2E-16	5.8E-16	2.0E-17	0.0E+00	1.5E-17
SO4	1.6E-17	2.9E-17	9.4E-16	3.6E-17	5.1E-16	9.7E-16	2.3E-15	5.8E-16	1.1E-15	2.2E-15
EC	1.9E-01	4.8E-03	3.4E-03	2.5E-01	7.8E-03	9.5E-04	3.9E-03	1.6E-04	0.0E+00	1.8E-03
Na	0.0E+00	0.0E+00	2.7E-04	6.7E-07	2.6E-05	1.7E-02	7.3E-05	8.1E-07	1.5E-01	5.1E-05
Al	0.0E+00	0.0E+00	2.8E-05	1.0E-06	5.5E-05	3.6E-03	4.7E-04	1.7E-03	0.0E+00	1.5E-06
Si	0.0E+00	3.7E-05	3.5E-05	2.1E-04	1.0E-02	1.6E-02	1.5E-02	2.5E-03	0.0E+00	3.5E-02
Cl	1.1E-05	0.0E+00	6.2E-04	2.5E-05	1.4E-04	1.0E-02	4.1E-03	1.2E-04	3.2E-01	7.0E-04
K	1.1E-06	0.0E+00	2.1E-05	1.4E-06	6.2E-06	5.5E-03	8.6E-03	6.4E-05	6.5E-03	2.0E-02
Ca	8.5E-05	8.9E-06	1.3E-04	1.4E-04	1.1E-03	1.0E-02	1.5E-03	4.8E-03	6.4E-03	9.0E-04
Ti	1.3E-07	0.0E+00	2.4E-05	2.7E-06	1.3E-04	1.3E-03	2.3E-05	6.6E-02	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	3.2E-06	5.9E-07	2.7E-05	7.2E-05	5.8E-06	0.0E+00	0.0E+00	0.0E+00
Fe	2.2E-05	3.1E-06	1.7E-05	5.2E-04	1.2E-02	2.4E-03	1.2E-03	1.0E-05	0.0E+00	1.1E-06
Cu	0.0E+00	0.0E+00	0.0E+00	3.2E-05	1.1E-03	2.8E-05	3.1E-04	5.1E-07	0.0E+00	0.0E+00
Zn	1.9E-04	1.5E-05	3.1E-06	2.3E-04	7.4E-04	1.5E-04	1.7E-03	3.6E-07	0.0E+00	4.4E-06
As	9.9E-08	0.0E+00	8.1E-09	1.0E-07	2.0E-07	1.4E-06	7.6E-07	0.0E+00	0.0E+00	2.7E-08
Se	0.0E+00	0.0E+00	8.6E-06	1.9E-08	3.8E-07	1.0E-06	6.3E-07	0.0E+00	0.0E+00	0.0E+00
Br	6.6E-07	0.0E+00	5.0E-06	7.0E-07	1.8E-06	8.0E-05	1.8E-06	3.2E-06	0.0E+00	2.6E-07
Sn	8.3E-07	0.0E+00	6.7E-08	3.7E-06	6.5E-05	2.6E-06	7.1E-07	0.0E+00	0.0E+00	1.4E-07
Sb	6.3E-07	6.7E-06	4.3E-06	6.5E-07	3.8E-06	4.9E-06	1.4E-06	1.6E-04	0.0E+00	0.0E+00
Ba	1.7E-05	0.0E+00	9.4E-07	6.2E-05	1.2E-03	2.4E-05	1.7E-05	5.4E-04	0.0E+00	0.0E+00
Pb	3.3E-08	1.0E-05	1.1E-08	2.5E-07	1.4E-05	1.0E-05	7.6E-07	9.9E-06	0.0E+00	8.2E-08

Table C.7. Summer (June, July, August) $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	1.8E-01	0.0E+00	3.9E-02	4.9E-02	4.8E-01	1.1E-01	6.5E-01	7.0E-02	4.8E-01
NO3	0.0E+00	1.1E-03	0.0E+00	2.0E-03	1.8E-03	5.6E-03	0.0E+00	4.6E-03	5.2E-03	3.3E-02
NH4	0.0E+00	0.0E+00	0.0E+00	2.4E-02	8.6E-04	6.2E-03	0.0E+00	0.0E+00	4.1E-04	4.2E-03
SO4	0.0E+00	3.0E-03	0.0E+00	1.2E-01	6.0E-03	8.4E-03	3.3E-01	2.6E-03	1.8E-01	1.2E-01
EC	0.0E+00	6.9E-01	0.0E+00	1.9E-02	1.6E-03	9.4E-02	4.7E-02	3.2E-02	1.3E-02	8.1E-02
Na	0.0E+00	0.0E+00	0.0E+00	1.1E-03	1.1E-03	3.9E-03	0.0E+00	3.1E-03	3.1E-02	2.0E-02
Al	0.0E+00	0.0E+00	0.0E+00	4.6E-02	6.2E-02	5.2E-04	0.0E+00	4.2E-04	3.6E-02	2.0E-03
Si	0.0E+00	0.0E+00	0.0E+00	1.1E-01	2.1E-01	9.7E-04	3.4E-03	7.8E-04	4.8E-02	2.6E-03
Cl	0.0E+00	2.0E-04	0.0E+00	5.7E-03	1.3E-03	2.2E-02	1.7E-03	8.7E-03	6.2E-02	3.5E-02
K	0.0E+00	3.6E-05	0.0E+00	5.3E-03	2.1E-02	1.6E-02	1.7E-04	2.9E-03	5.5E-02	1.6E-03
Ca	0.0E+00	5.7E-04	0.0E+00	5.2E-02	6.5E-02	2.2E-03	2.7E-04	5.1E-04	1.6E-02	1.1E-02
Ti	0.0E+00	3.8E-06	0.0E+00	2.5E-03	3.9E-03	3.2E-04	3.5E-04	2.7E-05	8.6E-04	9.1E-04
Mn	0.0E+00	0.0E+00	0.0E+00	1.7E-04	1.2E-03	6.0E-05	2.0E-05	6.4E-05	5.9E-03	1.8E-04
Fe	0.0E+00	2.5E-04	0.0E+00	3.0E-02	5.6E-02	4.2E-04	5.4E-04	7.8E-04	1.1E-01	8.3E-04
Cu	0.0E+00	0.0E+00	0.0E+00	1.5E-04	1.1E-04	1.8E-05	8.3E-05	1.2E-04	6.8E-03	0.0E+00
Zn	0.0E+00	7.2E-04	0.0E+00	5.7E-04	5.4E-04	1.1E-03	7.2E-04	2.6E-04	3.1E-03	0.0E+00
As	0.0E+00	2.9E-06	0.0E+00	1.2E-06	1.7E-05	5.9E-05	0.0E+00	2.8E-06	4.1E-04	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	3.1E-03	1.0E-06	2.9E-06	2.7E-05	9.6E-06	1.6E-04	5.6E-04
Br	0.0E+00	1.4E-05	0.0E+00	3.0E-04	1.2E-05	2.5E-04	0.0E+00	1.6E-04	1.5E-03	3.7E-04
Sn	0.0E+00	2.4E-05	0.0E+00	1.2E-04	2.7E-05	9.4E-05	0.0E+00	1.1E-05	3.6E-04	0.0E+00
Sb	0.0E+00	1.8E-05	0.0E+00	7.7E-06	3.1E-05	3.4E-06	1.3E-04	2.2E-05	7.8E-04	0.0E+00
Ba	0.0E+00	3.3E-04	0.0E+00	3.8E-03	7.5E-04	1.3E-05	0.0E+00	8.7E-04	2.7E-03	0.0E+00
Pb	0.0E+00	9.5E-07	0.0E+00	1.2E-04	2.1E-04	7.3E-05	0.0E+00	3.6E-04	2.1E-03	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	1.8E-01	4.8E-01	2.1E-01	1.8E-01	4.1E-01	1.1E-01	2.4E-01	2.7E-01	0.0E+00	4.2E-01
NO3	1.1E-03	7.0E-04	1.3E-02	1.2E-03	1.5E-03	9.2E-03	8.8E-03	3.5E-04	0.0E+00	7.8E-04
NH4	0.0E+00	0.0E+00	1.7E-03	1.7E-04	1.0E-02	7.2E-03	3.2E-02	7.5E-04	0.0E+00	6.1E-04
SO4	3.0E-03	5.0E-04	2.9E-01	3.9E-03	1.8E-02	1.4E-01	1.3E-01	2.8E-02	7.6E-02	4.0E-02
EC	7.2E-01	1.2E-01	8.8E-02	7.0E-01	1.4E-01	2.5E-02	9.5E-02	4.4E-03	0.0E+00	4.2E-02
Na	0.0E+00	0.0E+00	7.7E-03	1.9E-05	7.3E-04	4.3E-02	2.1E-03	2.3E-05	3.0E-01	1.2E-03
Al	0.0E+00	0.0E+00	7.9E-04	2.9E-05	1.3E-03	1.5E-02	7.8E-03	2.4E-02	0.0E+00	4.3E-05
Si	0.0E+00	1.2E-03	1.0E-03	1.4E-03	3.7E-02	5.6E-02	7.4E-02	2.6E-02	0.0E+00	9.4E-02
Cl	2.0E-04	0.0E+00	1.6E-02	3.3E-04	1.7E-03	1.7E-02	2.7E-02	3.0E-03	3.8E-01	4.2E-03
K	3.7E-05	0.0E+00	6.2E-04	4.4E-05	2.0E-04	2.7E-02	4.4E-02	1.9E-03	1.1E-02	7.3E-02
Ca	6.8E-04	2.9E-04	4.0E-03	8.7E-04	5.8E-03	3.6E-02	1.1E-02	2.9E-02	1.1E-02	3.9E-03
Ti	3.8E-06	0.0E+00	6.9E-04	5.6E-05	1.1E-03	4.7E-03	5.8E-04	1.7E-01	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	9.1E-05	1.6E-05	3.5E-04	5.4E-04	1.6E-04	0.0E+00	0.0E+00	0.0E+00
Fe	2.9E-04	9.6E-05	5.0E-04	2.8E-03	5.7E-02	1.2E-02	9.3E-03	3.2E-04	0.0E+00	3.5E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.9E-04	4.6E-03	2.5E-04	2.4E-03	1.5E-05	0.0E+00	0.0E+00
Zn	9.7E-04	3.1E-04	9.0E-05	1.1E-03	3.6E-03	8.8E-04	8.7E-03	1.0E-05	0.0E+00	8.7E-05
As	2.9E-06	0.0E+00	2.3E-07	3.0E-06	5.8E-06	3.4E-05	2.2E-05	0.0E+00	0.0E+00	7.8E-07
Se	0.0E+00	0.0E+00	2.5E-04	5.5E-07	1.1E-05	2.9E-05	1.8E-05	0.0E+00	0.0E+00	0.0E+00
Br	1.5E-05	0.0E+00	1.5E-04	1.5E-05	4.7E-05	4.8E-04	5.3E-05	9.5E-05	0.0E+00	7.8E-06
Sn	2.4E-05	0.0E+00	1.9E-06	1.1E-04	1.9E-03	7.5E-05	2.0E-05	0.0E+00	0.0E+00	3.9E-06
Sb	1.8E-05	1.9E-04	1.2E-04	1.9E-05	1.1E-04	1.4E-04	3.9E-05	4.5E-03	0.0E+00	0.0E+00
Ba	3.4E-04	0.0E+00	2.7E-05	8.0E-04	1.1E-02	6.0E-04	4.9E-04	1.2E-02	0.0E+00	0.0E+00
Pb	9.5E-07	2.9E-04	3.2E-07	7.2E-06	3.1E-04	2.2E-04	2.2E-05	2.9E-04	0.0E+00	2.3E-06

Table C.8. Summer (June, July, August) standard deviations for $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	7.2E-15	0.0E+00	9.7E-17	2.2E-15	7.0E-15	2.7E-15	3.3E-15	8.5E-16	8.3E-16
NO3	0.0E+00	4.1E-17	0.0E+00	6.6E-17	7.4E-17	1.3E-16	0.0E+00	1.6E-16	4.6E-17	3.0E-16
NH4	0.0E+00	0.0E+00	0.0E+00	4.1E-16	2.8E-17	5.6E-17	0.0E+00	0.0E+00	1.1E-17	6.9E-17
SO4	0.0E+00	0.0E+00	0.0E+00	6.5E-15	2.0E-16	2.7E-16	5.7E-15	2.3E-17	6.0E-15	3.3E-15
EC	0.0E+00	6.3E-02	0.0E+00	7.8E-04	6.4E-05	4.1E-03	1.9E-03	1.3E-03	5.6E-04	3.4E-03
Na	0.0E+00	0.0E+00	0.0E+00	5.0E-05	4.8E-05	1.5E-04	0.0E+00	1.4E-04	4.1E-03	2.6E-03
Al	0.0E+00	0.0E+00	0.0E+00	1.7E-02	1.9E-02	2.0E-05	0.0E+00	1.5E-05	5.9E-03	7.7E-05
Si	0.0E+00	0.0E+00	0.0E+00	3.3E-02	7.3E-02	3.7E-05	1.4E-04	2.6E-05	9.8E-03	1.1E-04
Cl	0.0E+00	7.6E-06	0.0E+00	1.5E-03	1.2E-04	1.9E-03	8.9E-05	1.6E-03	9.0E-02	1.1E-02
K	0.0E+00	1.4E-06	0.0E+00	1.1E-03	5.9E-03	1.7E-03	6.0E-06	2.6E-04	1.2E-02	1.0E-04
Ca	0.0E+00	2.8E-05	0.0E+00	1.9E-02	2.2E-02	1.0E-04	9.8E-06	1.9E-05	3.7E-03	1.7E-03
Ti	0.0E+00	1.5E-07	0.0E+00	4.1E-04	6.9E-04	1.2E-05	1.3E-05	1.0E-06	4.7E-05	4.3E-05
Mn	0.0E+00	0.0E+00	0.0E+00	1.5E-05	2.2E-04	2.3E-06	7.6E-07	2.4E-06	1.3E-03	8.4E-06
Fe	0.0E+00	9.5E-06	0.0E+00	7.4E-03	1.8E-02	1.6E-05	2.1E-05	3.7E-05	2.8E-02	4.1E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.4E-05	8.4E-06	6.9E-07	3.1E-06	5.5E-06	1.7E-03	0.0E+00
Zn	0.0E+00	8.6E-05	0.0E+00	9.6E-05	8.8E-05	6.1E-05	5.5E-05	1.3E-05	6.0E-04	0.0E+00
As	0.0E+00	1.1E-07	0.0E+00	4.6E-08	6.6E-07	2.3E-06	0.0E+00	1.1E-07	3.6E-05	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	7.8E-04	4.0E-08	1.1E-07	1.0E-06	3.7E-07	7.4E-06	3.0E-05
Br	0.0E+00	5.4E-07	0.0E+00	3.5E-05	4.6E-07	1.0E-05	0.0E+00	9.0E-06	2.4E-04	2.5E-05
Sn	0.0E+00	9.2E-07	0.0E+00	4.8E-06	1.0E-06	3.6E-06	0.0E+00	4.3E-07	1.4E-05	0.0E+00
Sb	0.0E+00	7.0E-07	0.0E+00	3.0E-07	1.2E-06	1.3E-07	4.9E-06	8.3E-07	3.0E-05	0.0E+00
Ba	0.0E+00	1.3E-05	0.0E+00	3.4E-04	3.7E-05	5.2E-07	0.0E+00	3.3E-05	1.3E-04	0.0E+00
Pb	0.0E+00	3.7E-08	0.0E+00	5.2E-06	1.0E-05	2.8E-06	0.0E+00	1.5E-05	2.3E-04	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	7.2E-15	1.5E-14	4.7E-16	6.0E-15	1.9E-14	5.9E-15	3.0E-15	8.3E-16	0.0E+00	1.8E-14
NO3	4.1E-17	1.2E-17	1.9E-16	1.1E-18	1.3E-17	1.7E-16	3.3E-16	5.2E-18	0.0E+00	2.2E-17
NH4	0.0E+00	0.0E+00	4.1E-17	4.9E-19	1.4E-16	2.9E-16	8.3E-16	1.5E-17	0.0E+00	8.7E-18
SO4	0.0E+00	2.7E-17	2.1E-15	6.4E-17	5.0E-16	8.9E-16	3.3E-15	5.7E-16	2.4E-16	2.1E-15
EC	2.0E-01	5.1E-03	3.8E-03	2.2E-01	7.3E-03	9.9E-04	4.3E-03	1.7E-04	0.0E+00	1.8E-03
Na	0.0E+00	0.0E+00	3.0E-04	7.3E-07	2.7E-05	1.9E-02	8.4E-05	8.9E-07	1.3E-01	5.0E-05
Al	0.0E+00	0.0E+00	3.1E-05	1.1E-06	4.8E-05	2.9E-03	4.0E-04	1.3E-03	0.0E+00	1.7E-06
Si	0.0E+00	3.9E-05	3.9E-05	1.7E-04	7.3E-03	1.3E-02	1.3E-02	2.0E-03	0.0E+00	2.4E-02
Cl	1.5E-05	0.0E+00	7.5E-04	3.5E-05	1.6E-04	1.2E-02	5.6E-03	1.4E-04	3.1E-01	4.0E-04
K	1.2E-06	0.0E+00	2.3E-05	1.6E-06	6.7E-06	5.4E-03	9.1E-03	7.9E-05	6.3E-03	1.9E-02
Ca	1.0E-04	9.7E-06	1.4E-04	1.6E-04	1.0E-03	1.1E-02	1.6E-03	5.3E-03	6.7E-03	5.6E-04
Ti	1.5E-07	0.0E+00	2.7E-05	2.9E-06	1.0E-04	1.2E-03	2.4E-05	7.2E-02	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	3.5E-06	7.5E-07	3.0E-05	8.6E-05	6.9E-06	0.0E+00	0.0E+00	0.0E+00
Fe	2.7E-05	3.3E-06	1.9E-05	5.5E-04	1.3E-02	2.6E-03	1.4E-03	1.1E-05	0.0E+00	1.3E-06
Cu	0.0E+00	0.0E+00	0.0E+00	3.6E-05	1.1E-03	3.3E-05	3.6E-04	5.6E-07	0.0E+00	0.0E+00
Zn	2.0E-04	1.9E-05	3.4E-06	2.4E-04	7.2E-04	1.5E-04	1.7E-03	4.0E-07	0.0E+00	3.7E-06
As	1.1E-07	0.0E+00	8.9E-09	1.1E-07	2.2E-07	1.4E-06	8.3E-07	0.0E+00	0.0E+00	3.0E-08
Se	0.0E+00	0.0E+00	9.4E-06	2.1E-08	4.2E-07	1.1E-06	7.0E-07	0.0E+00	0.0E+00	0.0E+00
Br	7.4E-07	0.0E+00	5.5E-06	7.7E-07	1.8E-06	7.5E-05	2.0E-06	3.5E-06	0.0E+00	3.0E-07
Sn	9.1E-07	0.0E+00	7.4E-08	4.1E-06	6.8E-05	2.9E-06	7.8E-07	0.0E+00	0.0E+00	1.5E-07
Sb	7.0E-07	7.3E-06	4.7E-06	7.2E-07	4.1E-06	5.4E-06	1.5E-06	1.7E-04	0.0E+00	0.0E+00
Ba	1.8E-05	0.0E+00	1.0E-06	6.2E-05	1.1E-03	2.7E-05	1.9E-05	6.2E-04	0.0E+00	0.0E+00
Pb	3.7E-08	1.2E-05	1.2E-08	2.7E-07	1.4E-05	1.1E-05	8.4E-07	1.1E-05	0.0E+00	9.0E-08

Table C.9. Fall (September, October, November) $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	1.8E-01	0.0E+00	3.9E-02	4.9E-02	4.8E-01	1.1E-01	6.5E-01	7.0E-02	4.8E-01
NO3	0.0E+00	1.1E-03	0.0E+00	2.0E-03	1.8E-03	5.6E-03	0.0E+00	4.6E-03	5.2E-03	3.3E-02
NH4	0.0E+00	0.0E+00	0.0E+00	2.4E-02	8.6E-04	6.2E-03	0.0E+00	0.0E+00	4.1E-04	4.2E-03
SO4	0.0E+00	2.9E-03	0.0E+00	1.2E-01	6.0E-03	8.4E-03	3.3E-01	2.6E-03	1.8E-01	1.2E-01
EC	0.0E+00	7.0E-01	0.0E+00	2.0E-02	1.7E-03	9.5E-02	4.7E-02	3.3E-02	1.3E-02	8.1E-02
Na	0.0E+00	0.0E+00	0.0E+00	1.1E-03	1.1E-03	3.9E-03	0.0E+00	3.1E-03	3.1E-02	2.0E-02
Al	0.0E+00	0.0E+00	0.0E+00	4.7E-02	6.1E-02	5.3E-04	0.0E+00	4.2E-04	3.6E-02	2.0E-03
Si	0.0E+00	0.0E+00	0.0E+00	9.6E-02	1.9E-01	9.7E-04	3.4E-03	7.8E-04	4.6E-02	2.6E-03
Cl	0.0E+00	2.0E-04	0.0E+00	5.5E-03	1.3E-03	2.2E-02	1.7E-03	8.4E-03	4.9E-02	3.3E-02
K	0.0E+00	3.6E-05	0.0E+00	5.3E-03	2.1E-02	1.7E-02	1.7E-04	2.9E-03	5.5E-02	1.6E-03
Ca	0.0E+00	5.7E-04	0.0E+00	5.1E-02	6.4E-02	2.3E-03	2.7E-04	5.1E-04	1.6E-02	1.1E-02
Ti	0.0E+00	3.8E-06	0.0E+00	2.3E-03	3.7E-03	3.2E-04	3.5E-04	2.7E-05	8.5E-04	9.0E-04
Mn	0.0E+00	0.0E+00	0.0E+00	1.8E-04	1.2E-03	6.1E-05	2.0E-05	6.4E-05	6.0E-03	1.8E-04
Fe	0.0E+00	2.5E-04	0.0E+00	3.0E-02	5.7E-02	4.2E-04	5.4E-04	7.7E-04	1.1E-01	8.4E-04
Cu	0.0E+00	0.0E+00	0.0E+00	1.5E-04	1.1E-04	1.8E-05	8.4E-05	1.2E-04	6.8E-03	0.0E+00
Zn	0.0E+00	7.1E-04	0.0E+00	5.6E-04	5.3E-04	1.1E-03	7.3E-04	2.6E-04	3.0E-03	0.0E+00
As	0.0E+00	2.9E-06	0.0E+00	1.2E-06	1.7E-05	5.9E-05	0.0E+00	2.8E-06	4.0E-04	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	3.0E-03	1.1E-06	2.9E-06	2.7E-05	9.6E-06	1.6E-04	5.6E-04
Br	0.0E+00	1.4E-05	0.0E+00	3.1E-04	1.2E-05	2.6E-04	0.0E+00	1.6E-04	1.6E-03	3.8E-04
Sn	0.0E+00	2.4E-05	0.0E+00	1.2E-04	2.7E-05	9.4E-05	0.0E+00	1.1E-05	3.6E-04	0.0E+00
Sb	0.0E+00	1.8E-05	0.0E+00	7.8E-06	3.1E-05	3.4E-06	1.3E-04	2.2E-05	7.8E-04	0.0E+00
Ba	0.0E+00	3.3E-04	0.0E+00	3.8E-03	7.5E-04	1.3E-05	0.0E+00	8.8E-04	2.7E-03	0.0E+00
Pb	0.0E+00	9.5E-07	0.0E+00	1.2E-04	2.1E-04	7.3E-05	0.0E+00	3.7E-04	2.2E-03	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	1.8E-01	4.8E-01	2.1E-01	1.8E-01	4.1E-01	1.1E-01	2.4E-01	2.7E-01	0.0E+00	4.2E-01
NO3	1.1E-03	7.0E-04	1.3E-02	1.2E-03	1.5E-03	9.2E-03	8.8E-03	3.5E-04	0.0E+00	7.8E-04
NH4	0.0E+00	0.0E+00	1.7E-03	1.7E-04	1.0E-02	7.2E-03	3.2E-02	7.5E-04	0.0E+00	6.1E-04
SO4	2.9E-03	5.0E-04	2.9E-01	3.9E-03	1.8E-02	1.4E-01	1.3E-01	2.8E-02	7.6E-02	4.0E-02
EC	7.7E-01	1.2E-01	8.8E-02	7.5E-01	1.4E-01	2.5E-02	9.6E-02	4.4E-03	0.0E+00	4.3E-02
Na	0.0E+00	0.0E+00	7.7E-03	1.9E-05	7.3E-04	4.1E-02	2.1E-03	2.3E-05	3.1E-01	1.2E-03
Al	0.0E+00	0.0E+00	7.9E-04	3.0E-05	1.3E-03	1.5E-02	7.8E-03	2.4E-02	0.0E+00	4.3E-05
Si	0.0E+00	1.2E-03	1.0E-03	1.4E-03	3.6E-02	5.0E-02	7.2E-02	2.5E-02	0.0E+00	1.1E-01
Cl	2.0E-04	0.0E+00	1.6E-02	3.2E-04	1.7E-03	1.5E-02	2.6E-02	3.0E-03	3.1E-01	4.2E-03
K	3.7E-05	0.0E+00	6.2E-04	4.4E-05	2.0E-04	2.6E-02	4.4E-02	1.9E-03	1.2E-02	8.9E-02
Ca	6.7E-04	2.9E-04	4.0E-03	8.6E-04	6.0E-03	3.7E-02	1.1E-02	2.9E-02	1.2E-02	4.6E-03
Ti	3.8E-06	0.0E+00	6.9E-04	5.6E-05	1.1E-03	4.2E-03	5.8E-04	1.5E-01	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	9.1E-05	1.6E-05	3.6E-04	5.4E-04	1.7E-04	0.0E+00	0.0E+00	0.0E+00
Fe	2.9E-04	9.6E-05	5.0E-04	2.8E-03	5.8E-02	1.2E-02	9.2E-03	3.2E-04	0.0E+00	3.5E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.9E-04	4.7E-03	2.4E-04	2.4E-03	1.5E-05	0.0E+00	0.0E+00
Zn	9.5E-04	3.0E-04	9.1E-05	1.1E-03	3.6E-03	8.3E-04	8.6E-03	1.0E-05	0.0E+00	8.9E-05
As	2.9E-06	0.0E+00	2.3E-07	3.0E-06	5.8E-06	3.3E-05	2.2E-05	0.0E+00	0.0E+00	7.8E-07
Se	0.0E+00	0.0E+00	2.5E-04	5.5E-07	1.1E-05	2.9E-05	1.8E-05	0.0E+00	0.0E+00	0.0E+00
Br	1.5E-05	0.0E+00	1.5E-04	1.5E-05	4.7E-05	4.9E-04	5.3E-05	9.5E-05	0.0E+00	7.8E-06
Sn	2.4E-05	0.0E+00	1.9E-06	1.1E-04	1.9E-03	7.6E-05	2.0E-05	0.0E+00	0.0E+00	3.9E-06
Sb	1.8E-05	1.9E-04	1.2E-04	1.9E-05	1.1E-04	1.4E-04	3.9E-05	4.5E-03	0.0E+00	0.0E+00
Ba	3.4E-04	0.0E+00	2.7E-05	7.9E-04	1.0E-02	6.0E-04	4.9E-04	1.2E-02	0.0E+00	0.0E+00
Pb	9.6E-07	2.9E-04	3.2E-07	7.2E-06	3.1E-04	2.2E-04	2.2E-05	2.9E-04	0.0E+00	2.3E-06

Table C.10. Fall (September, October, November) standard deviations for $PM_{2.5}$ source profiles averaged over all available CSN monitors during the 2006 study. Values represent the mass fraction of each species emitted by each source.

	ag	air	biog	coal	dust	fire	foil	meat	metal	ng
OC	0.0E+00	6.7E-15	0.0E+00	4.8E-16	2.1E-15	4.7E-15	2.6E-15	3.3E-15	1.7E-15	8.3E-16
NO3	0.0E+00	3.7E-17	0.0E+00	6.3E-17	6.2E-17	1.3E-16	0.0E+00	1.5E-16	2.1E-17	3.0E-16
NH4	0.0E+00	0.0E+00	0.0E+00	4.0E-16	2.3E-17	1.0E-17	0.0E+00	0.0E+00	9.8E-18	8.3E-17
SO4	0.0E+00	2.4E-17	0.0E+00	5.7E-15	1.3E-16	2.3E-16	5.6E-15	1.0E-17	4.0E-15	3.1E-15
EC	0.0E+00	6.4E-02	0.0E+00	7.4E-04	6.0E-05	3.9E-03	1.7E-03	1.2E-03	5.1E-04	3.2E-03
Na	0.0E+00	0.0E+00	0.0E+00	4.3E-05	4.1E-05	1.6E-04	0.0E+00	1.2E-04	3.1E-03	1.4E-03
Al	0.0E+00	0.0E+00	0.0E+00	1.5E-02	1.7E-02	1.9E-05	0.0E+00	1.4E-05	5.6E-03	7.2E-05
Si	0.0E+00	0.0E+00	0.0E+00	4.7E-02	8.1E-02	3.6E-05	1.4E-04	2.6E-05	1.0E-02	1.2E-04
Cl	0.0E+00	7.1E-06	0.0E+00	1.5E-03	9.0E-05	2.9E-03	9.0E-05	1.5E-03	9.2E-02	1.2E-02
K	0.0E+00	1.3E-06	0.0E+00	1.0E-03	5.5E-03	2.6E-03	5.5E-06	2.5E-04	1.2E-02	1.0E-04
Ca	0.0E+00	2.5E-05	0.0E+00	2.2E-02	2.2E-02	1.5E-04	9.0E-06	2.1E-05	3.7E-03	1.9E-03
Ti	0.0E+00	1.4E-07	0.0E+00	2.3E-04	3.7E-04	1.2E-05	1.2E-05	9.6E-07	3.4E-05	3.6E-05
Mn	0.0E+00	0.0E+00	0.0E+00	1.5E-05	2.2E-04	2.3E-06	7.1E-07	2.3E-06	1.4E-03	8.2E-06
Fe	0.0E+00	8.4E-06	0.0E+00	7.3E-03	1.7E-02	1.7E-05	1.9E-05	3.4E-05	2.9E-02	3.7E-05
Cu	0.0E+00	0.0E+00	0.0E+00	1.3E-05	7.1E-06	6.5E-07	3.0E-06	5.9E-06	1.6E-03	0.0E+00
Zn	0.0E+00	6.8E-05	0.0E+00	7.7E-05	6.6E-05	8.6E-05	5.6E-05	1.3E-05	5.1E-04	0.0E+00
As	0.0E+00	1.0E-07	0.0E+00	4.3E-08	6.1E-07	2.1E-06	0.0E+00	1.0E-07	2.3E-05	0.0E+00
Se	0.0E+00	0.0E+00	0.0E+00	5.6E-04	3.8E-08	1.0E-07	9.7E-07	3.4E-07	6.8E-06	2.7E-05
Br	0.0E+00	5.0E-07	0.0E+00	3.6E-05	4.3E-07	1.2E-05	0.0E+00	8.9E-06	2.5E-04	2.6E-05
Sn	0.0E+00	8.6E-07	0.0E+00	4.5E-06	9.6E-07	3.4E-06	0.0E+00	4.0E-07	1.3E-05	0.0E+00
Sb	0.0E+00	6.5E-07	0.0E+00	2.8E-07	1.1E-06	1.2E-07	4.6E-06	7.8E-07	2.8E-05	0.0E+00
Ba	0.0E+00	1.2E-05	0.0E+00	2.2E-04	3.2E-05	4.8E-07	0.0E+00	3.2E-05	1.2E-04	0.0E+00
Pb	0.0E+00	3.4E-08	0.0E+00	5.3E-06	9.2E-06	2.7E-06	0.0E+00	1.6E-05	2.5E-04	0.0E+00

	nrd	nrg	nro	ord	org	ot	otc	slv	ss	wood
OC	6.7E-15	1.5E-14	4.7E-16	4.0E-15	1.6E-14	5.0E-15	1.8E-15	8.3E-16	0.0E+00	1.4E-14
NO3	3.7E-17	5.0E-18	1.9E-16	2.0E-17	1.5E-18	1.6E-16	3.2E-16	5.1E-18	0.0E+00	1.6E-17
NH4	0.0E+00	0.0E+00	2.8E-17	4.9E-19	1.0E-17	2.4E-16	7.8E-16	8.1E-18	0.0E+00	6.5E-19
SO4	2.4E-17	2.4E-17	5.9E-15	6.3E-17	4.8E-16	2.3E-15	3.1E-15	5.6E-16	9.6E-16	1.9E-15
EC	1.5E-01	4.8E-03	3.5E-03	1.7E-01	8.6E-03	9.6E-04	4.0E-03	1.6E-04	0.0E+00	1.7E-03
Na	0.0E+00	0.0E+00	2.8E-04	6.9E-07	2.6E-05	9.1E-03	7.6E-05	8.3E-07	1.4E-01	4.5E-05
Al	0.0E+00	0.0E+00	2.9E-05	1.0E-06	4.8E-05	2.5E-03	3.6E-04	1.4E-03	0.0E+00	1.5E-06
Si	0.0E+00	3.8E-05	3.6E-05	1.8E-04	9.9E-03	1.7E-02	1.4E-02	2.0E-03	0.0E+00	4.0E-02
Cl	1.3E-05	0.0E+00	6.9E-04	3.3E-05	1.7E-04	1.2E-02	4.7E-03	1.2E-04	3.5E-01	7.4E-04
K	1.2E-06	0.0E+00	2.2E-05	1.5E-06	6.5E-06	5.4E-03	8.7E-03	6.7E-05	5.1E-03	2.0E-02
Ca	1.0E-04	9.1E-06	1.3E-04	1.6E-04	1.2E-03	1.3E-02	1.5E-03	4.9E-03	5.5E-03	8.8E-04
Ti	1.4E-07	0.0E+00	2.5E-05	2.2E-06	7.1E-05	6.8E-04	2.1E-05	6.1E-02	0.0E+00	0.0E+00
Mn	0.0E+00	0.0E+00	3.3E-06	7.4E-07	3.8E-05	8.1E-05	6.6E-06	0.0E+00	0.0E+00	0.0E+00
Fe	2.4E-05	3.2E-06	1.8E-05	5.5E-04	1.3E-02	2.6E-03	1.2E-03	1.1E-05	0.0E+00	1.1E-06
Cu	0.0E+00	0.0E+00	0.0E+00	3.6E-05	1.2E-03	2.8E-05	3.5E-04	5.2E-07	0.0E+00	0.0E+00
Zn	1.9E-04	1.3E-05	3.2E-06	2.2E-04	7.2E-04	1.2E-04	1.6E-03	3.8E-07	0.0E+00	3.8E-06
As	1.0E-07	0.0E+00	8.3E-09	1.0E-07	2.1E-07	1.2E-06	7.8E-07	0.0E+00	0.0E+00	2.8E-08
Se	0.0E+00	0.0E+00	8.8E-06	2.0E-08	4.0E-07	1.1E-06	6.5E-07	0.0E+00	0.0E+00	0.0E+00
Br	6.5E-07	0.0E+00	5.2E-06	7.2E-07	1.8E-06	6.8E-05	1.9E-06	3.3E-06	0.0E+00	2.7E-07
Sn	8.6E-07	0.0E+00	6.9E-08	3.8E-06	6.4E-05	2.7E-06	7.3E-07	0.0E+00	0.0E+00	1.4E-07
Sb	6.5E-07	6.9E-06	4.4E-06	6.7E-07	3.9E-06	5.0E-06	1.4E-06	1.6E-04	0.0E+00	0.0E+00
Ba	1.6E-05	0.0E+00	9.7E-07	4.9E-05	8.3E-04	2.6E-05	1.8E-05	4.7E-04	0.0E+00	0.0E+00
Pb	3.4E-08	1.2E-05	1.1E-08	2.5E-07	1.8E-05	1.2E-05	7.8E-07	1.0E-05	0.0E+00	8.4E-08

Table C.11. Cosine similarities (unitless) for optimized, seasonally-averaged regional profiles for coal combustion in comparison to the Reff et al. (2009) reference profile.

region number	winter	spring	summer	fall
1	1.00	1.00	1.00	1.00
2	0.99	0.99	0.98	0.99
3	1.00	0.97	0.98	1.00
4	0.99	0.98	0.99	0.99
5	1.00	0.99	0.99	1.00
6	0.99	0.99	0.99	0.99
7	0.98	0.97	0.98	0.99
8	0.99	0.99	1.00	0.99
9	0.99	0.99	0.99	0.99
10	0.99	0.98	0.97	
US	0.99	0.98	0.99	0.99

Table C.12. Cosine similarities (unitless) for optimized, seasonally-averaged regional profiles for metals processing in comparison to the Reff et al. (2009) reference profile.

region number	winter	spring	summer	fall
1	0.99	1.00	1.00	1.00
2	0.98	0.99	0.99	0.98
3	0.95	0.97	0.97	0.95
4	0.96	0.97	0.97	0.95
5	0.96	0.96	0.93	0.93
6	0.99	0.99	0.99	0.99
7	0.98	0.98	0.96	0.96
8	1.00	1.00	1.00	0.99
9	0.99	0.99	0.99	0.97
10	1.00	1.00	1.00	-
US	0.97	0.98	0.97	0.96

Table C.13. Cosine similarities (unitless) for optimized, seasonally-averaged regional profiles for sea salt in comparison to the Reff et al. (2009) reference profile.

region number	winter	spring	summer	fall
1	0.75	0.92	0.77	0.74
2	0.61	0.78	0.61	0.67
3	0.95	0.98	1.00	0.94
4	0.88	0.96	1.00	0.97
5	1.00	1.00	1.00	1.00
6	0.95	0.96	1.00	0.99
7	0.99	1.00	1.00	0.99
8	1.00	1.00	1.00	1.00
9	0.99	0.87	0.52	0.67
10	0.81	1.00	0.94	-
US	0.96	0.98	0.99	0.96

Table C.14. Mean (μ) and standard deviation (σ) of observations, original concentrations (SHreference), and revised (SHrevised) concentrations of metals for all available CSN site-days in 2006 (ug/m3).

	observations		$SH_{reference}$		$SH_{revised}$	
	μ	σ	μ	σ	μ	σ
Na	6.9E-02	1.1E-01	4.7E-02	3.8E-02	5.0E-02	4.4E-02
Al	3.8E-02	6.8E-02	2.9E-02	2.6E-02	3.3E-02	3.4E-02
Si	1.1E-01	1.8E-01	1.2E-01	1.2E-01	1.6E-01	1.3E-01
Cl	2.2E-02	8.0E-02	7.8E-02	7.3E-02	4.5E-02	6.9E-02
K	9.0E-02	2.4E-01	6.1E-02	7.5E-02	8.0E-02	6.1E-02
Ca	5.6E-02	7.4E-02	4.0E-02	2.8E-02	5.7E-02	4.1E-02
Ti	4.4E-03	7.8E-03	8.4E-03	6.2E-03	8.3E-03	7.6E-03
Mn	3.8E-03	2.2E-02	1.4E-03	1.3E-03	1.6E-03	1.7E-03
Fe	9.7E-02	1.4E-01	4.0E-02	3.2E-02	6.0E-02	4.7E-02
Cu	5.2E-03	1.0E-02	2.6E-03	2.2E-03	3.3E-03	3.1E-03
Zn	1.8E-02	5.7E-02	6.5E-03	5.0E-03	9.0E-03	7.1E-03
As	1.5E-03	1.7E-03	1.1E-04	1.1E-04	1.1E-04	1.2E-04
Se	1.7E-03	1.6E-03	6.9E-04	9.7E-04	7.5E-04	1.1E-03
Br	3.6E-03	4.0E-03	1.0E-03	7.3E-04	1.1E-03	8.4E-04
Sn	1.1E-02	4.7E-03	1.1E-03	9.2E-04	1.1E-03	9.2E-04
Sb	1.6E-02	6.3E-03	5.2E-04	3.8E-04	5.0E-04	3.8E-04
Ba	1.2E-02	3.2E-02	9.6E-03	7.9E-03	9.6E-03	8.0E-03
Pb	5.5E-03	1.2E-02	1.4E-03	1.1E-03	1.4E-03	1.1E-03

Table C.15. Correlations and mean normalized bias for original and revised concentrations of metals for all available CSN site-days in 2006 (ug/m3).

	correlation		mean normalized bias (MNB)	
	$SH_{reference}$ vs. obs	$SH_{revised}$ vs. obs	$SH_{reference}$ vs. obs	$SH_{revised}$ vs. obs
Na	0.27	0.27	0.49	0.50
Al	0.23	0.35	0.97	0.81
Si	0.21	0.40	1.58	1.38
Cl	0.16	0.35	12.72	4.10
K	0.06	0.09	0.06	0.40
Ca	0.34	0.41	0.58	0.80
Ti	0.22	0.39	1.98	1.57
Mn	0.18	0.19	-0.19	-0.14
Fe	0.36	0.35	-0.32	-0.08
Cu	0.21	0.24	0.09	0.24
Zn	0.15	0.11	-0.17	0.07
As	0.22	0.25	-0.89	-0.89
Se	0.06	0.12	-0.42	-0.40
Br	0.24	0.27	-0.49	-0.47
Sn	0.03	0.03	-0.88	-0.88
Sb	0.06	0.06	-0.96	-0.96
Ba	0.01	0.03	0.63	0.59
Pb	0.19	0.21	-0.51	-0.52

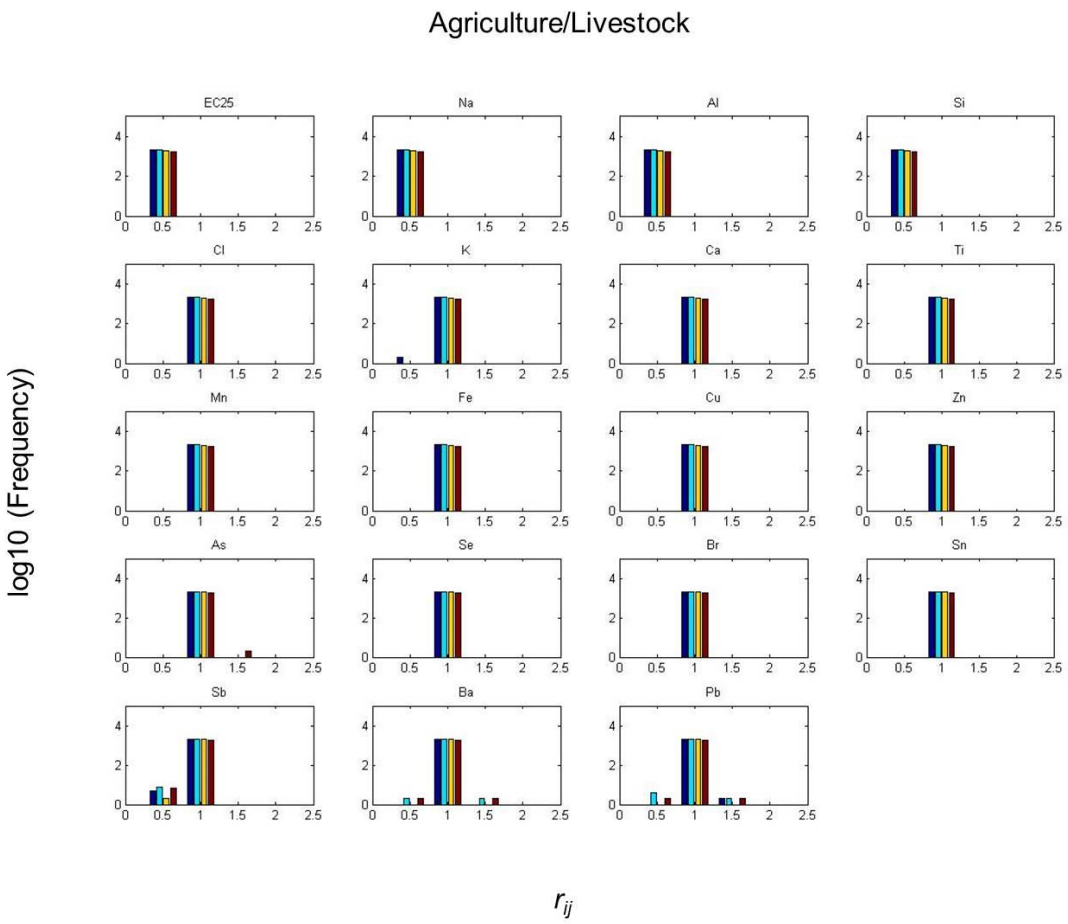

Figure C.1. Distribution of seasonally-stratified ratios rij for optimized agriculture/livestock source profiles for all available CSN monitors in 2006.

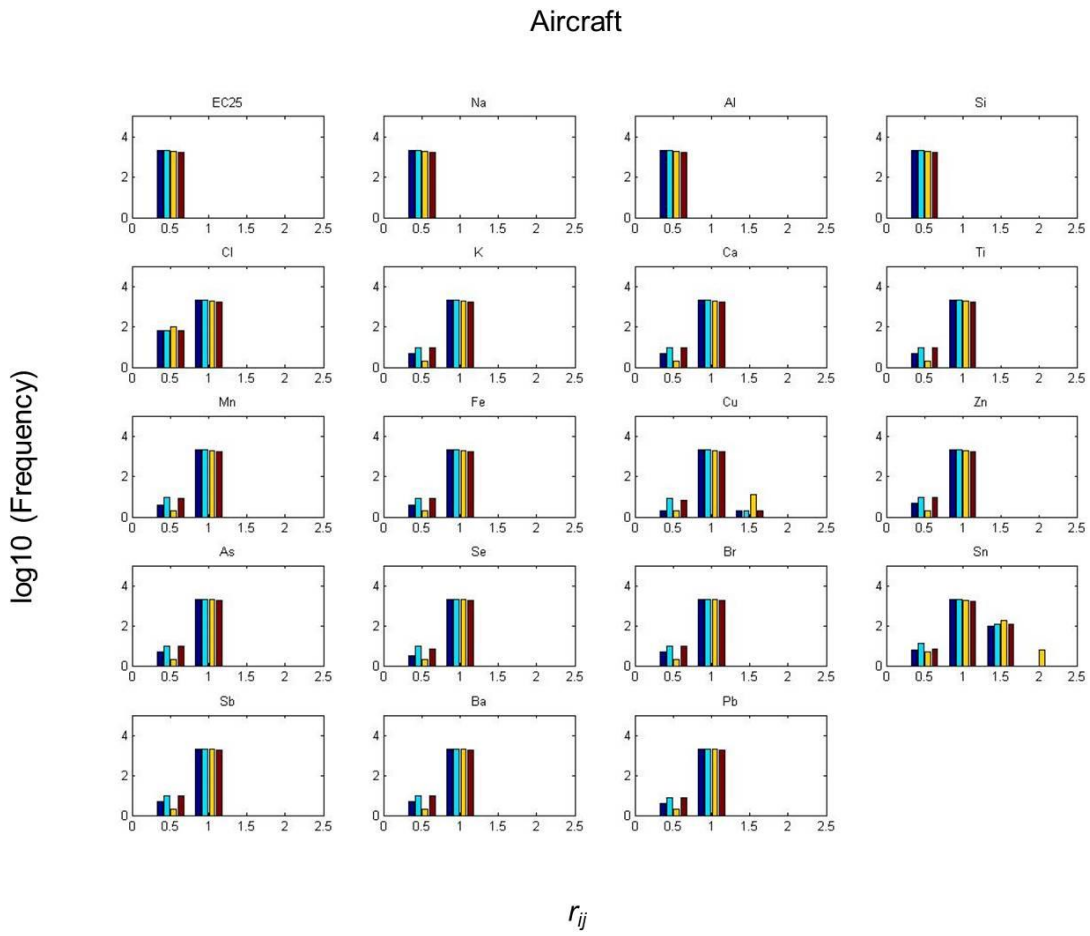

Figure C.2. Distribution of seasonally-stratified ratios rij for optimized aircraft source profiles for all available CSN monitors in 2006.

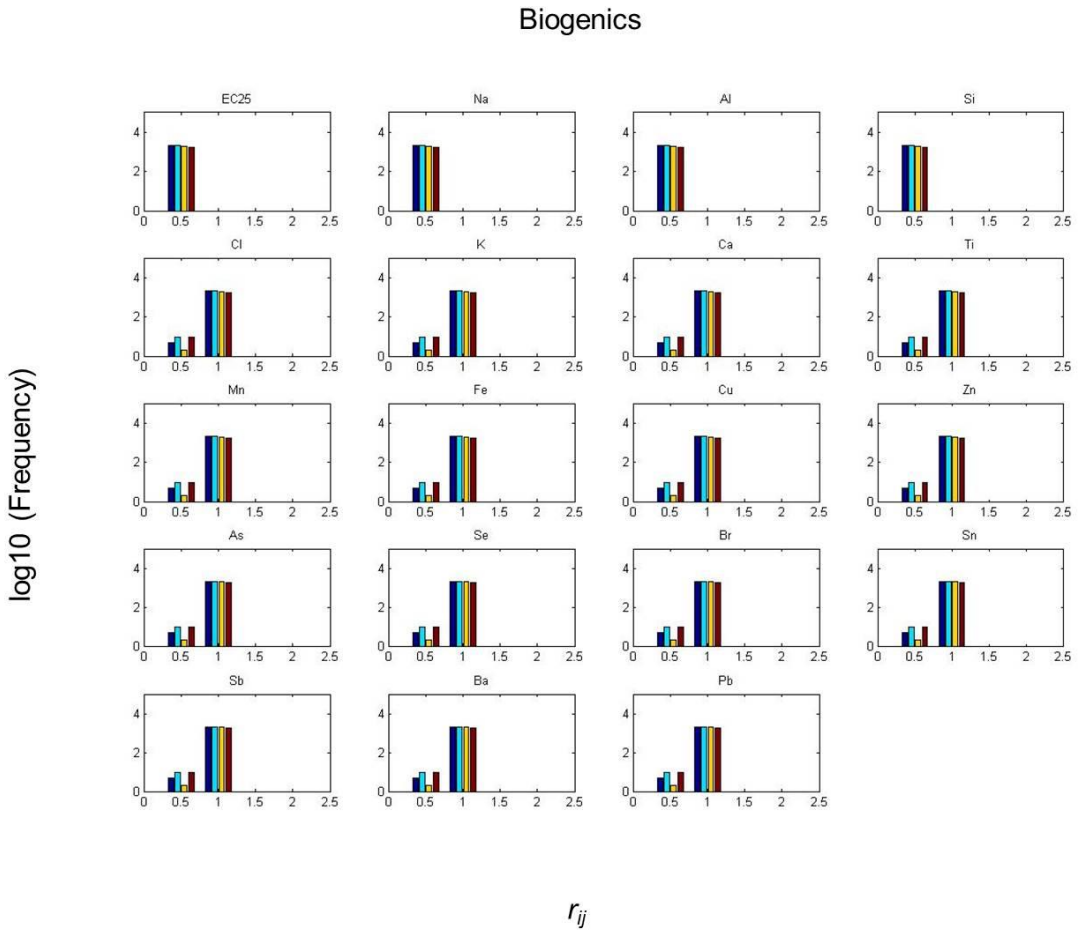

Figure C.3. Distribution of seasonally-stratified ratios rij for optimized biogenic source profiles for all available CSN monitors in 2006.

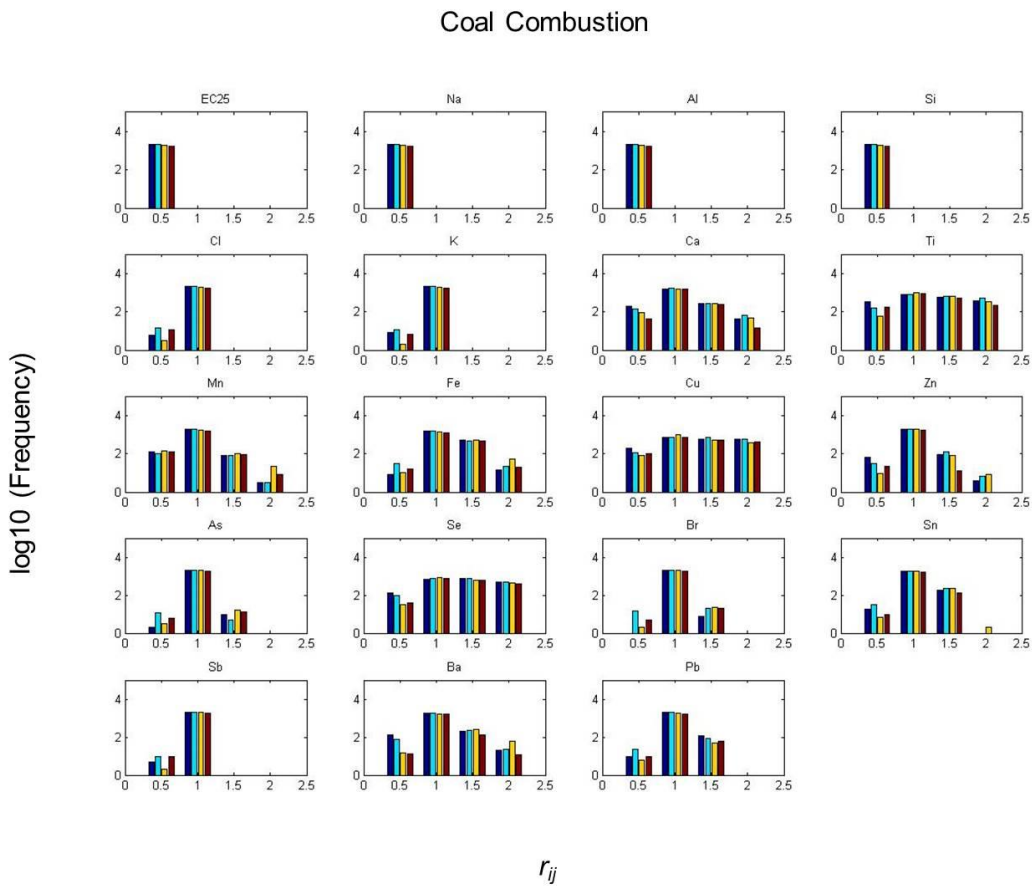

Figure C.4. Distribution of seasonally-stratified ratios rij for optimized coal combustion source profiles for all available CSN monitors in 2006.

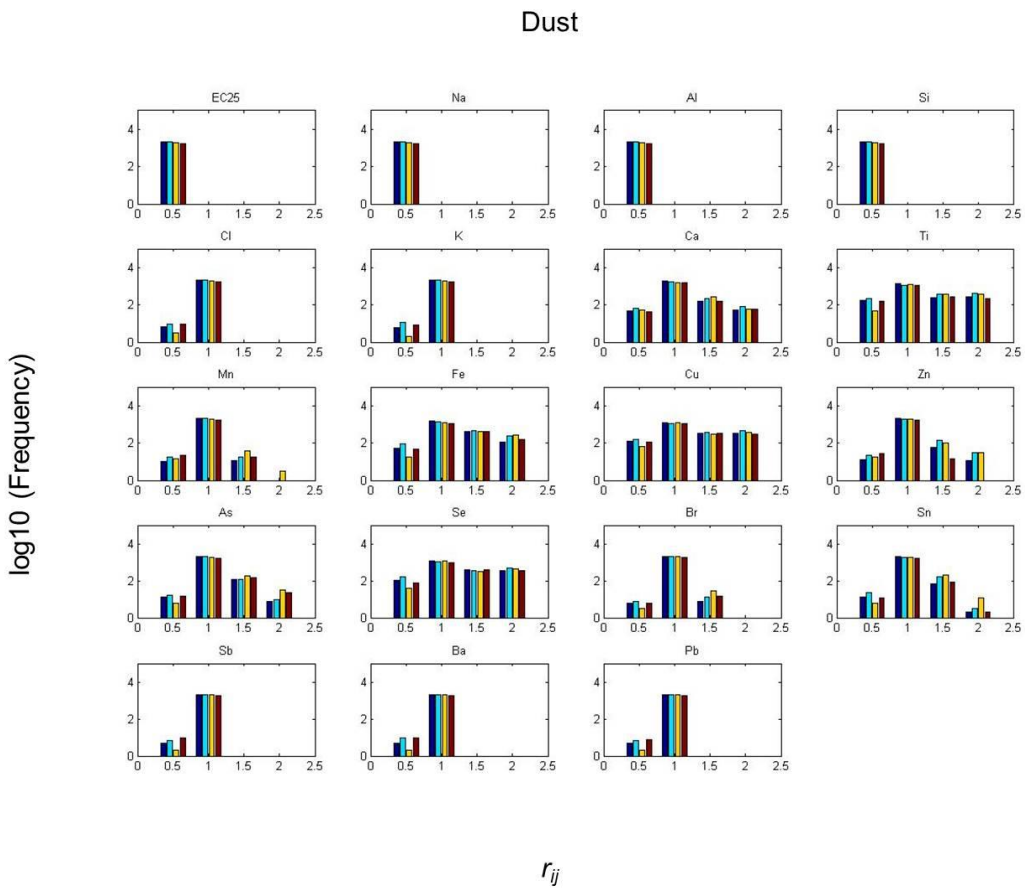

Figure C.5. Distribution of seasonally-stratified ratios rij for optimized dust source profiles for all available CSN monitors in 2006.

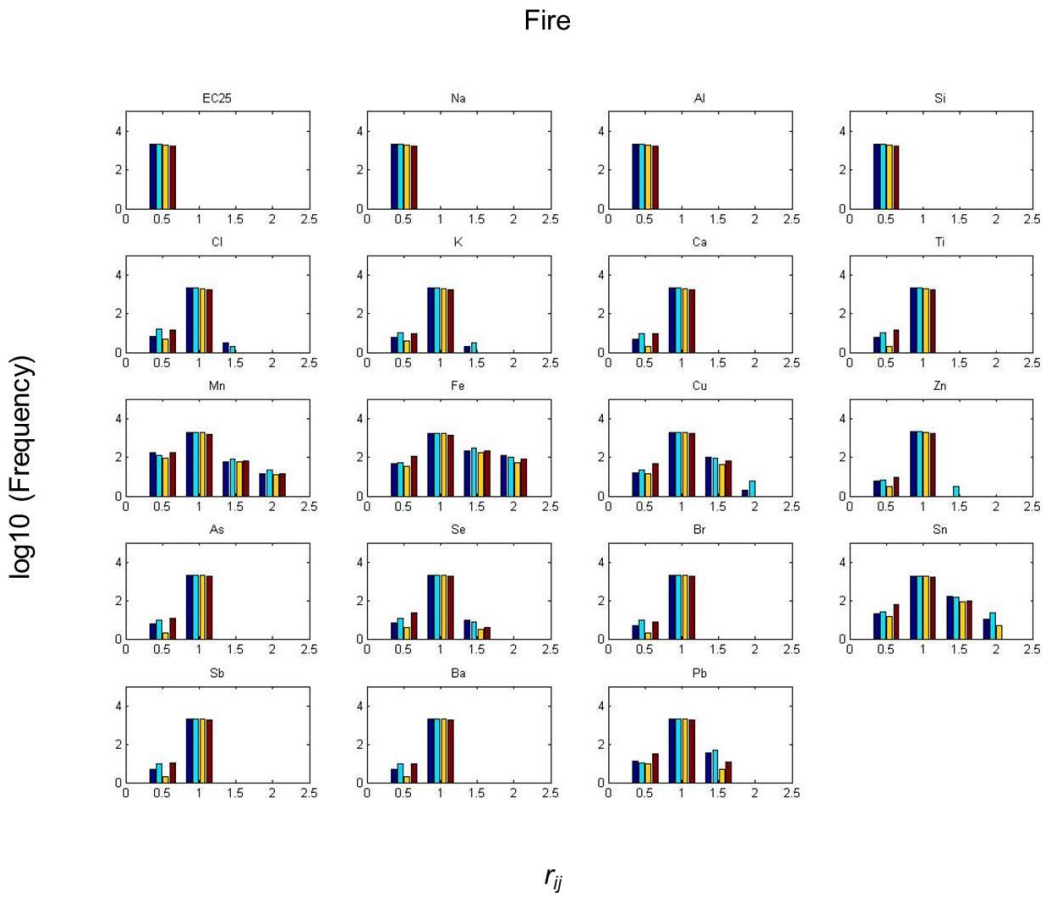

Figure C.6. Distribution of seasonally-stratified ratios rij for optimized fire source profiles for all available CSN monitors in 2006.

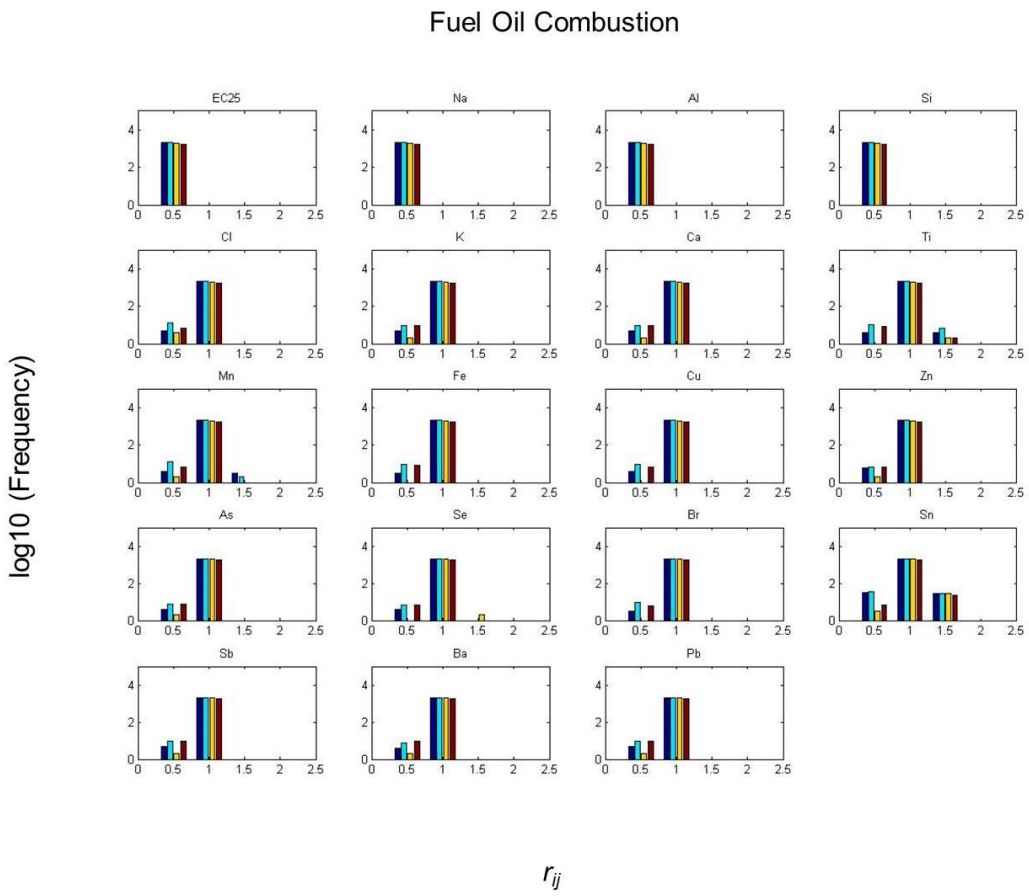

Figure C.7. Distribution of seasonally-stratified ratios rij for optimized fuel oil combustion source profiles for all available CSN monitors in 2006.

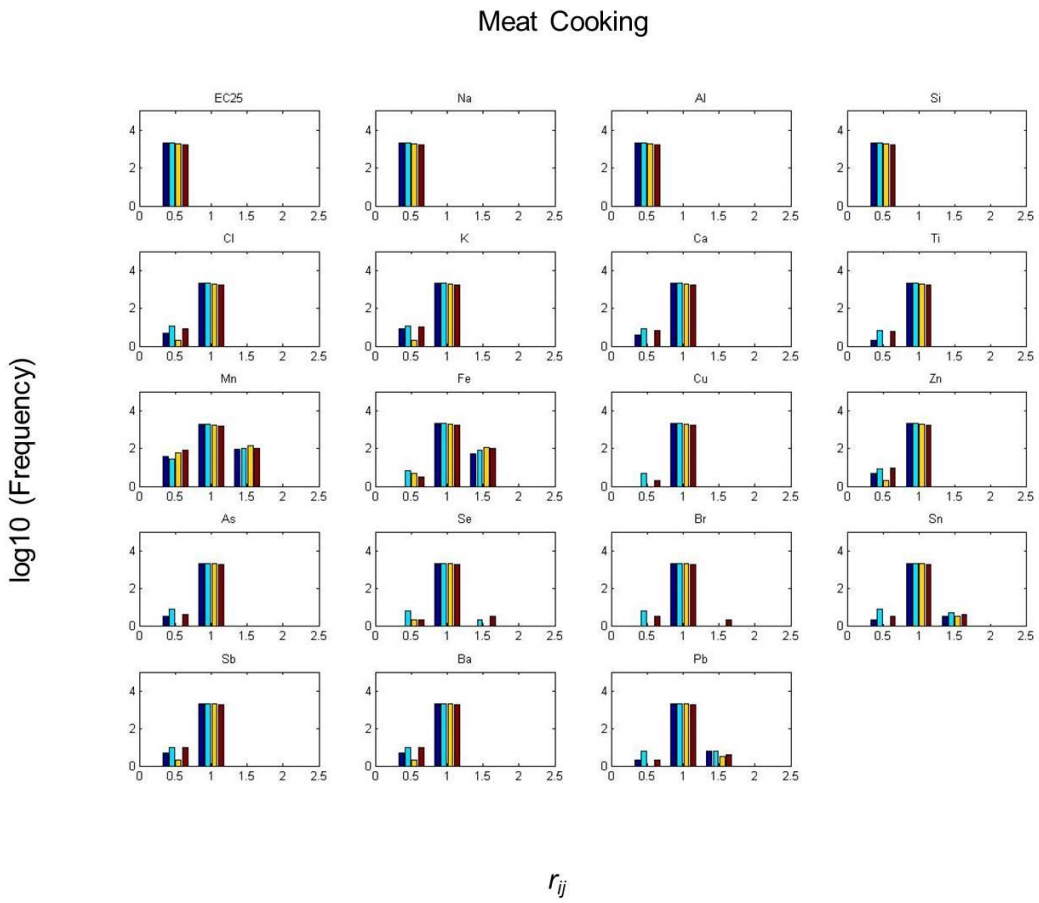

Figure C.8. Distribution of seasonally-stratified ratios rij for optimized meat cooking source profiles for all available CSN monitors in 2006.

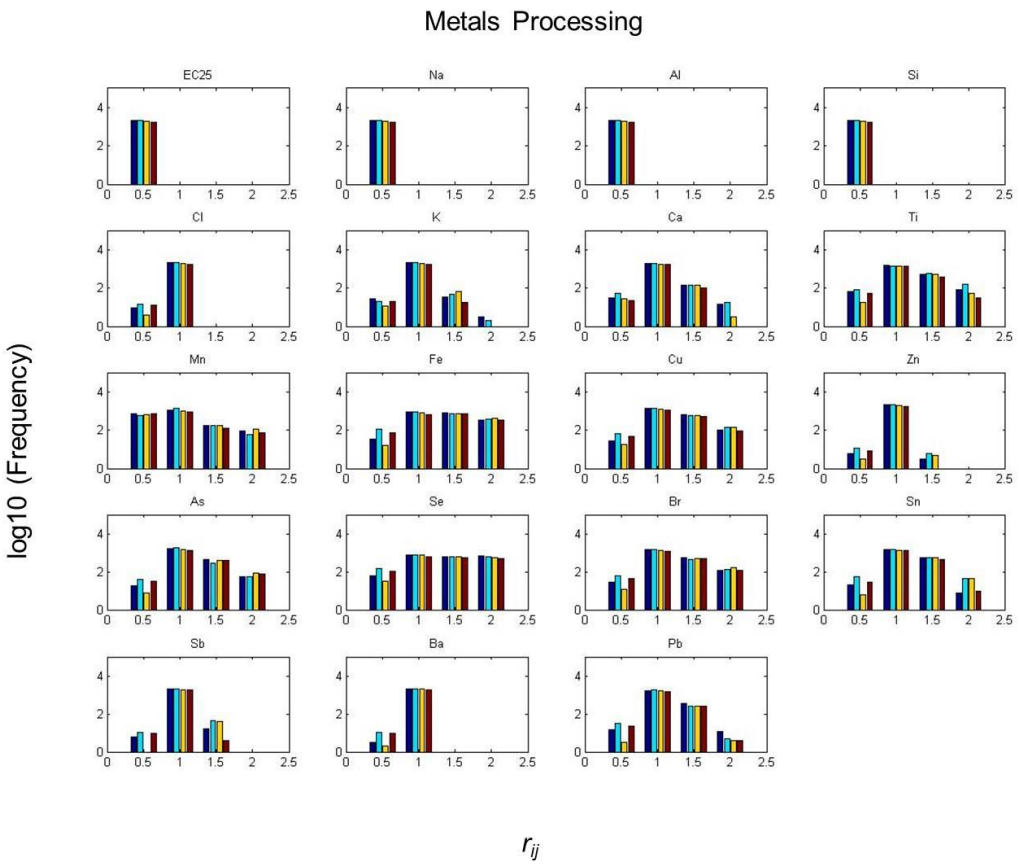

Figure C.9. Distribution of seasonally-stratified ratios rij for optimized metals processing source profiles for all available CSN monitors in 2006.

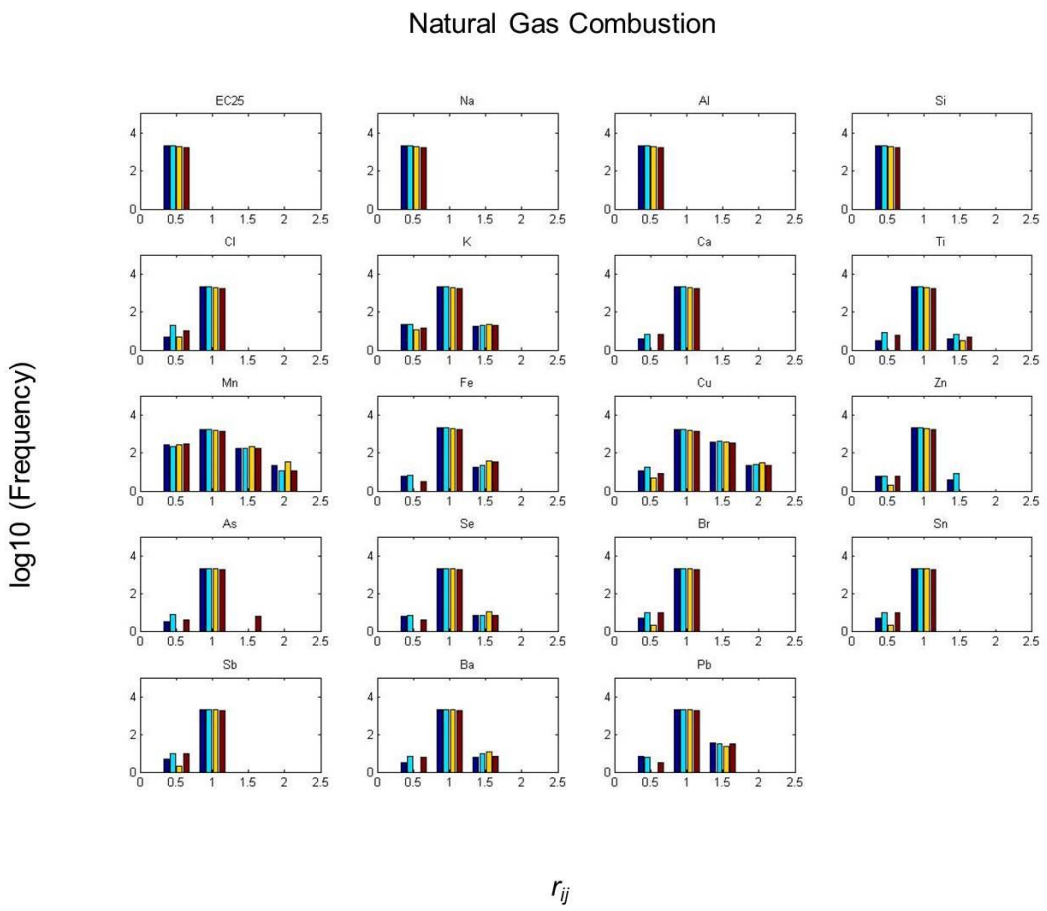

Figure C.10. Distribution of seasonally-stratified ratios rij for optimized natural gas combustion source profiles for all available CSN monitors in 2006.

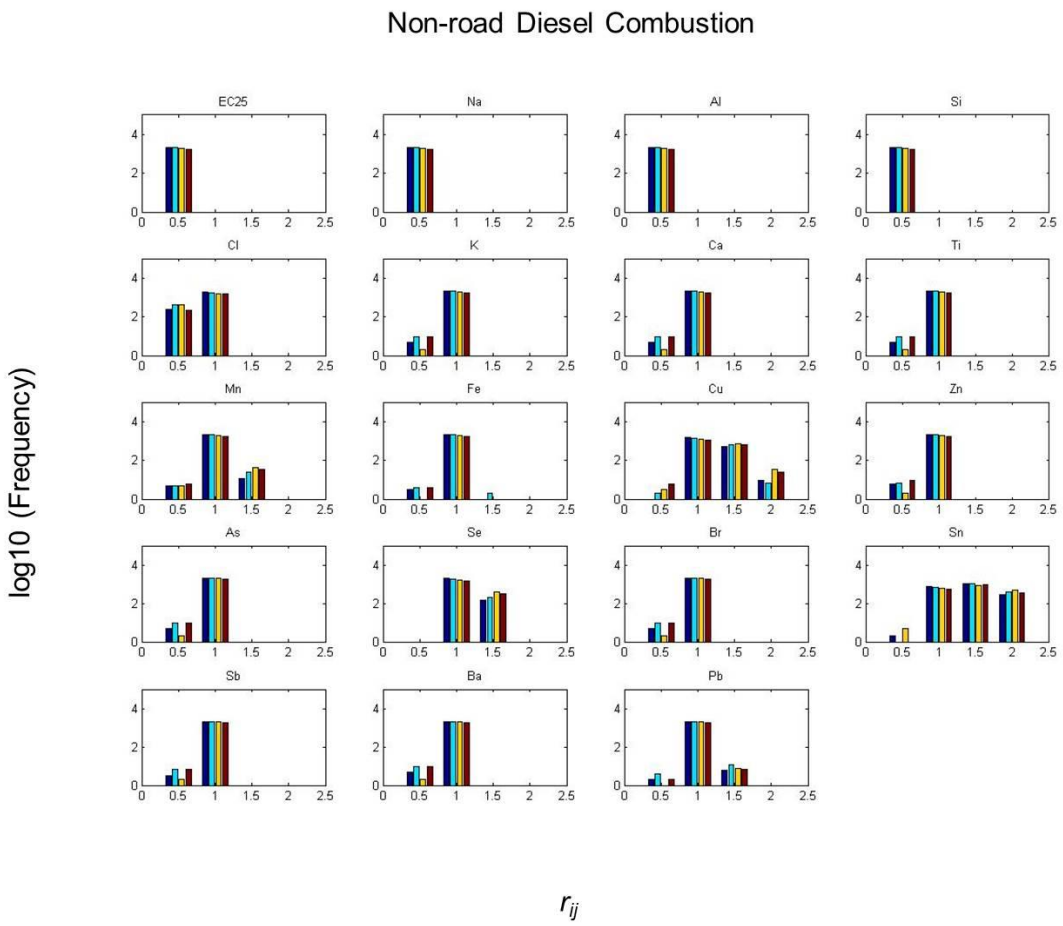

Figure C.11. Distribution of seasonally-stratified ratios rij for optimized non-road diesel combustion source profiles for all available CSN monitors in 2006.

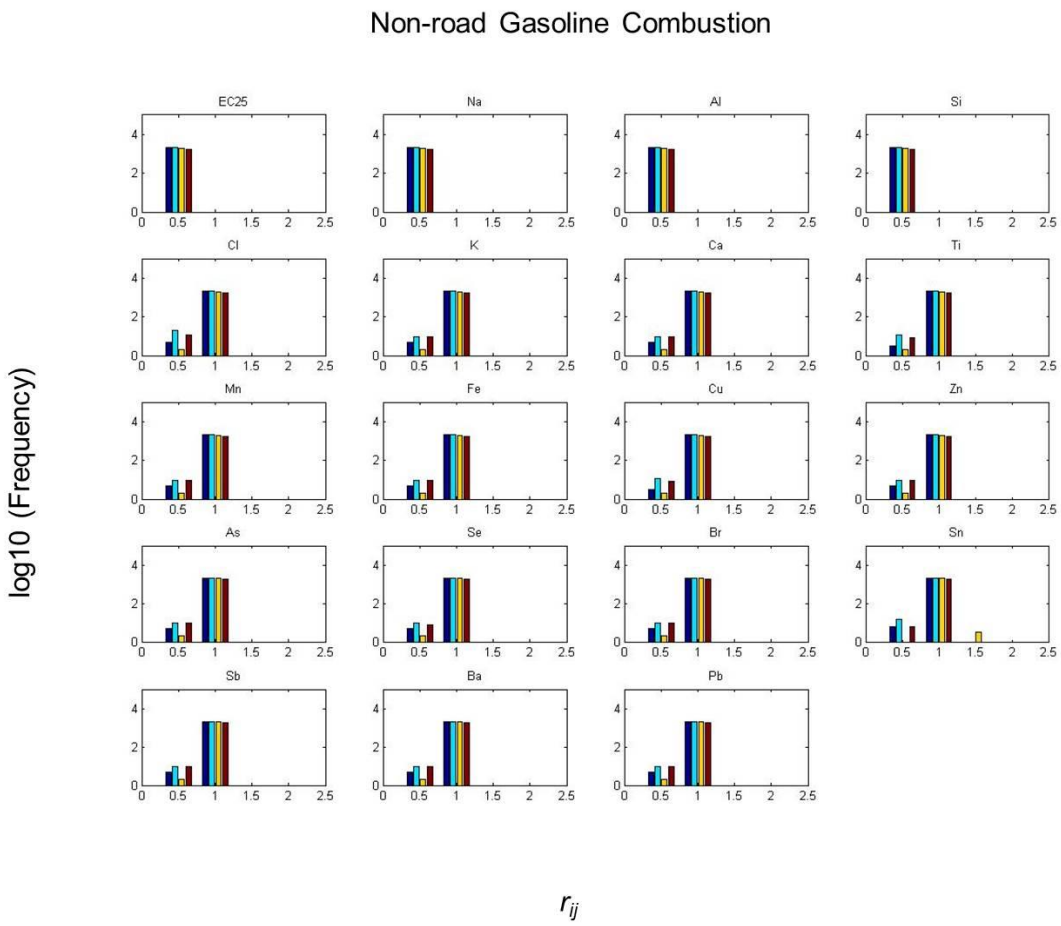

Figure C.12. Distribution of seasonally-stratified ratios rij for optimized non-road gasoline combustion source profiles for all available CSN monitors in 2006.

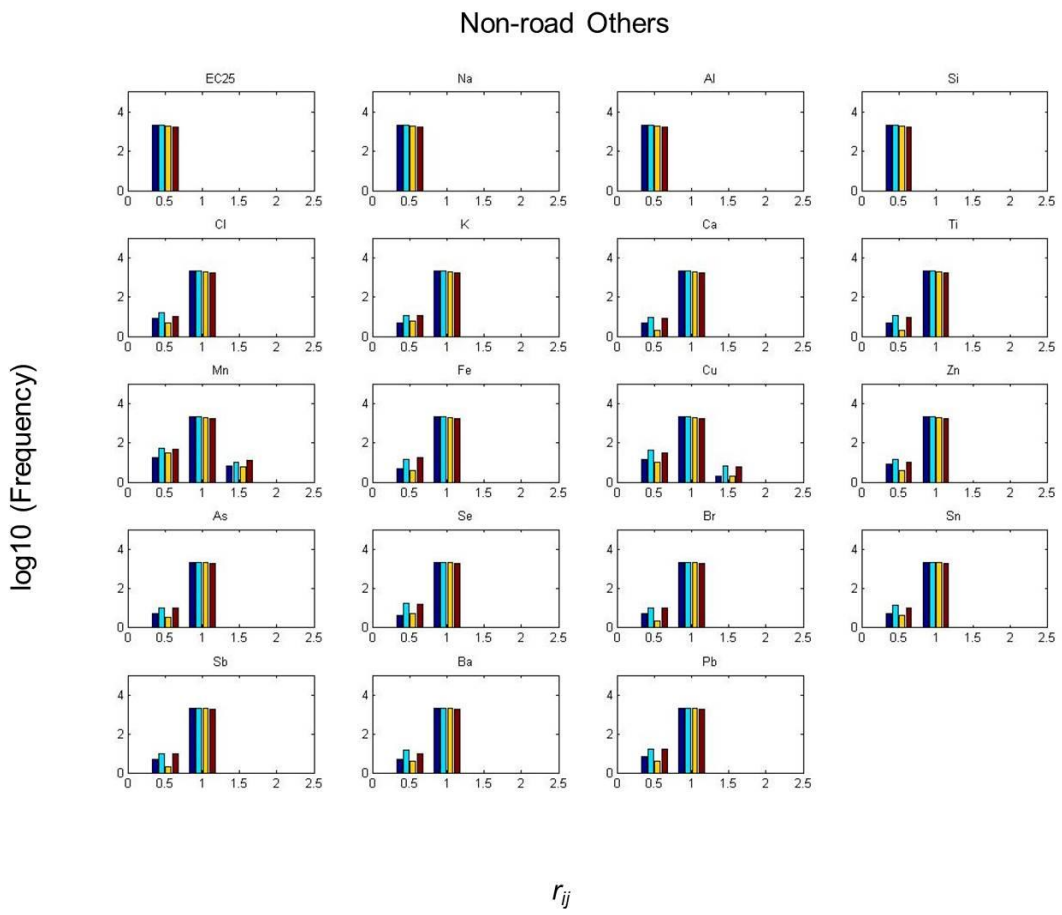

Figure C.13. Distribution of seasonally-stratified ratios rij for optimized non-road others source profiles for all available CSN monitors in 2006.

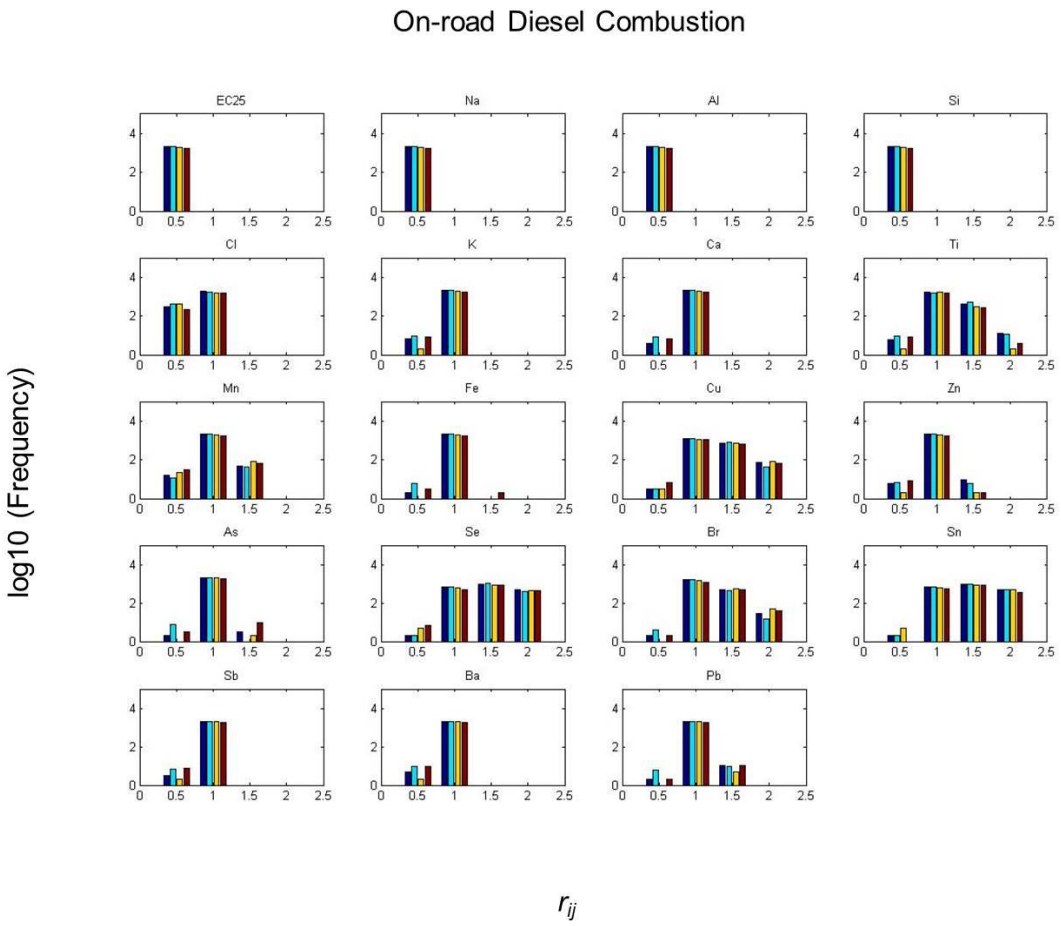

Figure C.14. Distribution of seasonally-stratified ratios rij for optimized on-road diesel combustion source profiles for all available CSN monitors in 2006.

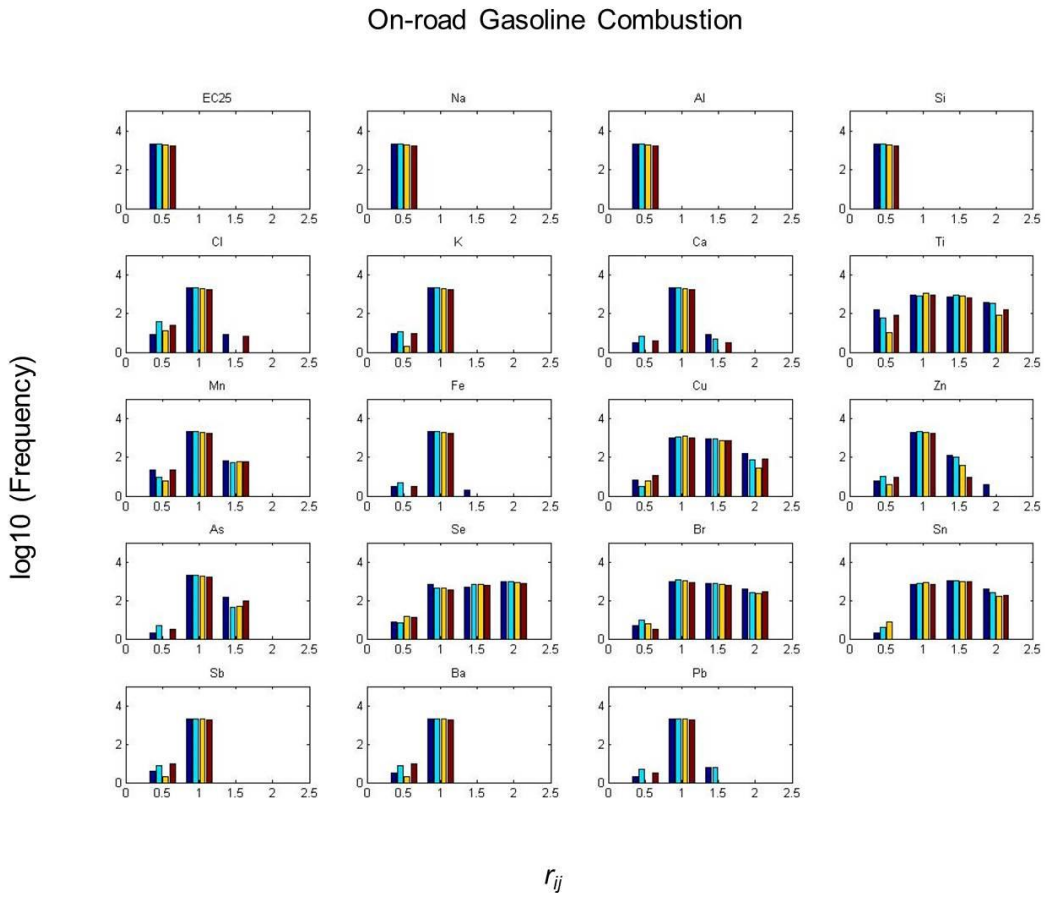

Figure C.15. Distribution of seasonally-stratified ratios rij for optimized on-road gasoline combustion source profiles for all available CSN monitors in 2006.

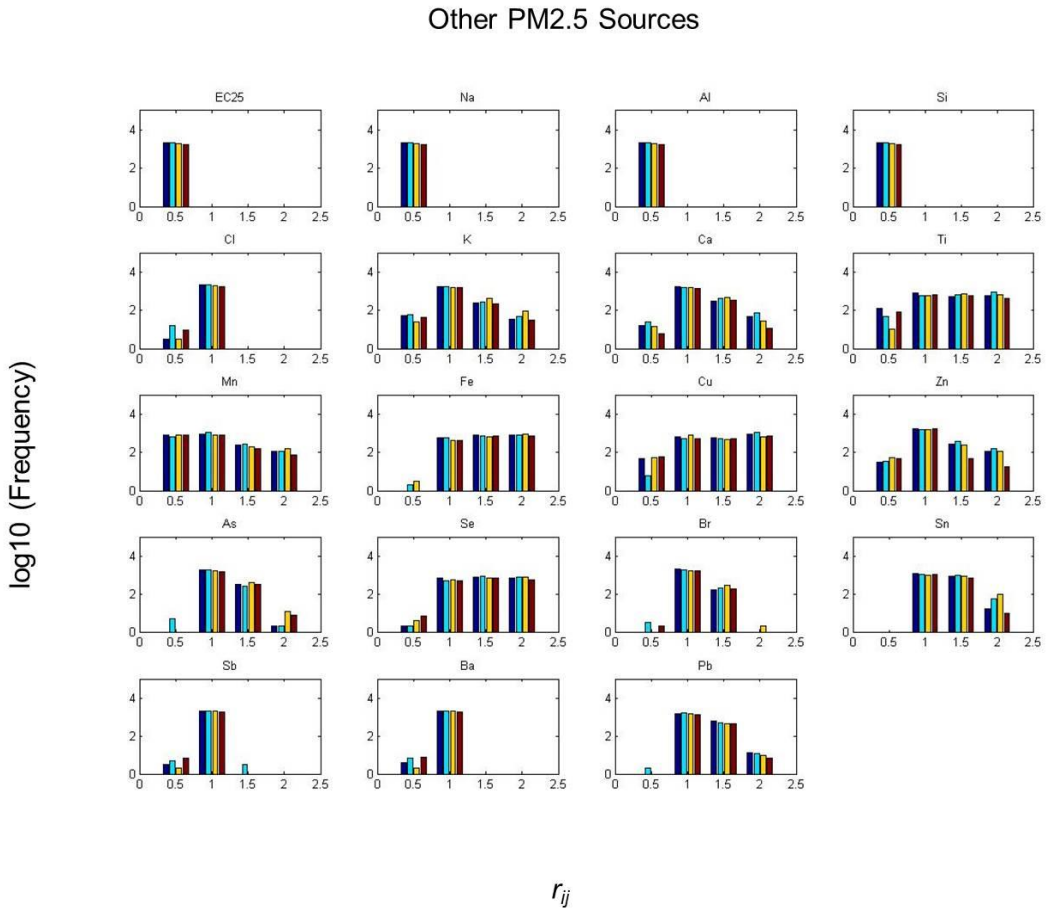

Figure C.16. Distribution of seasonally-stratified ratios rij for optimized other $PM_{2.5}$ source profiles for all available CSN monitors in 2006.

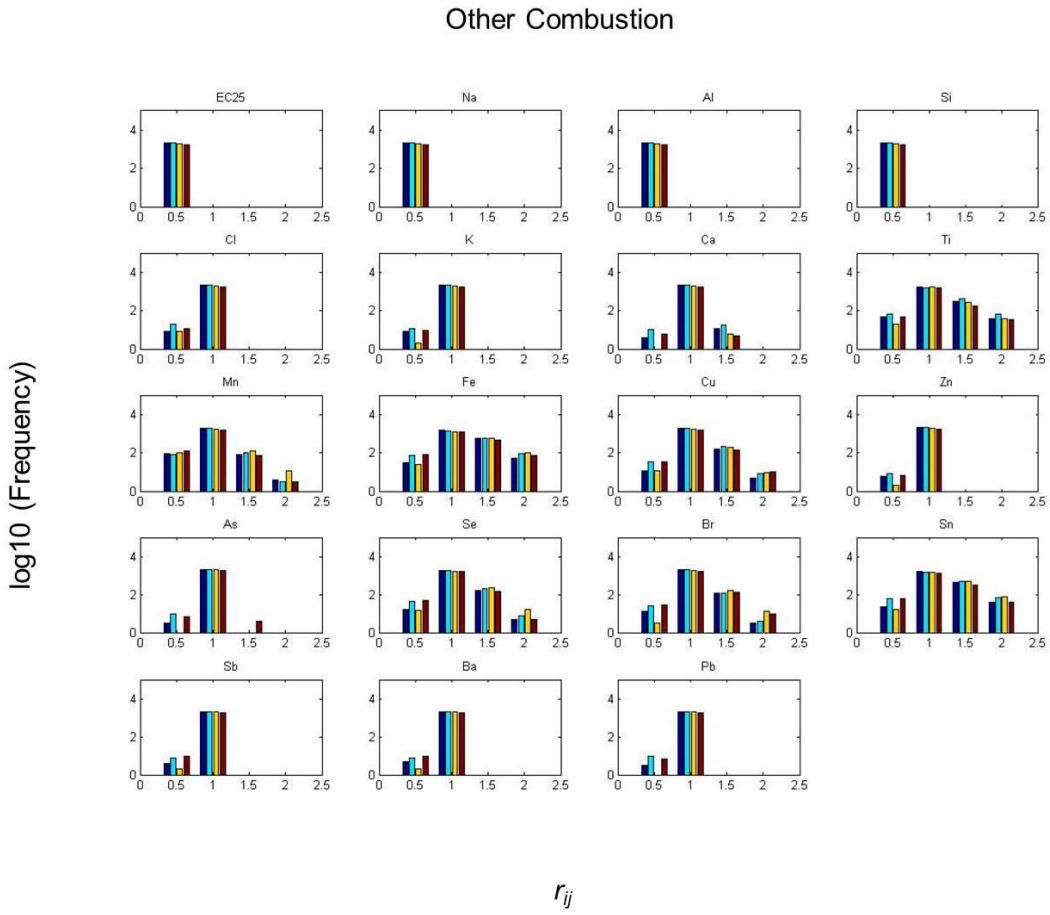

Figure C.17. Distribution of seasonally-stratified ratios rij for optimized other combustion source profiles for all available CSN monitors in 2006.

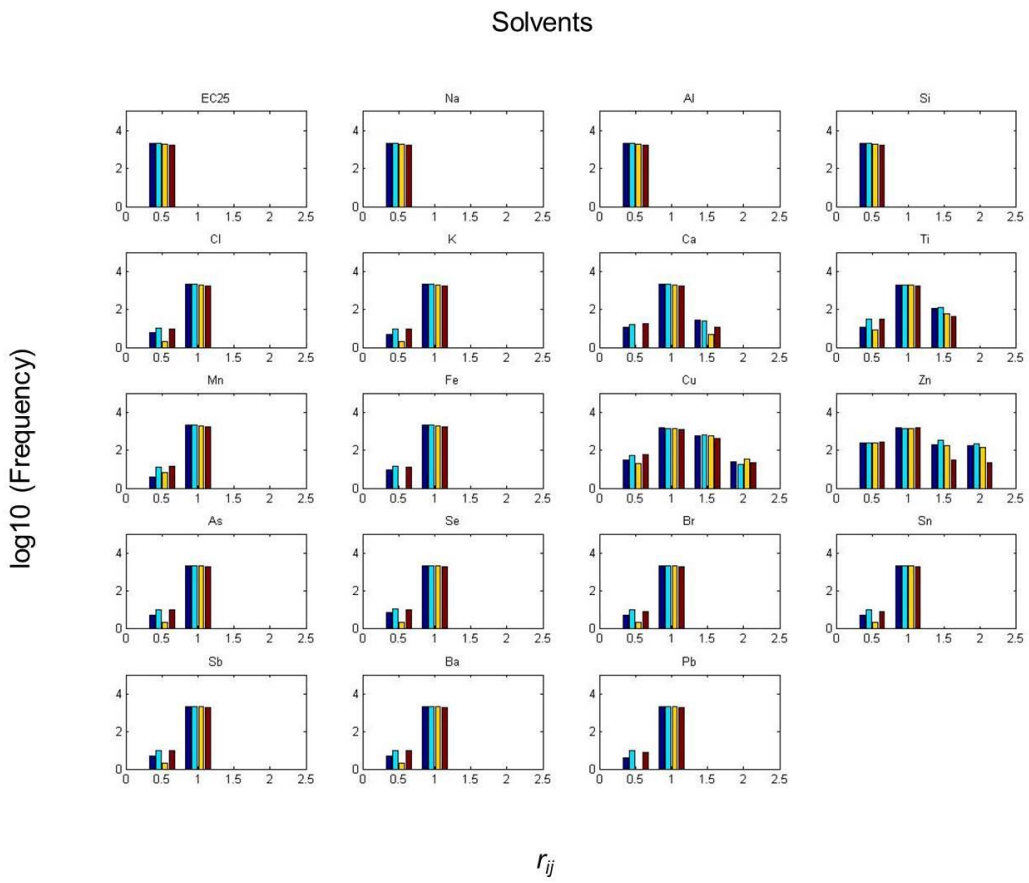

Figure C.18. Distribution of seasonally-stratified ratios rij for optimized solvents source profiles for all available CSN monitors in 2006.

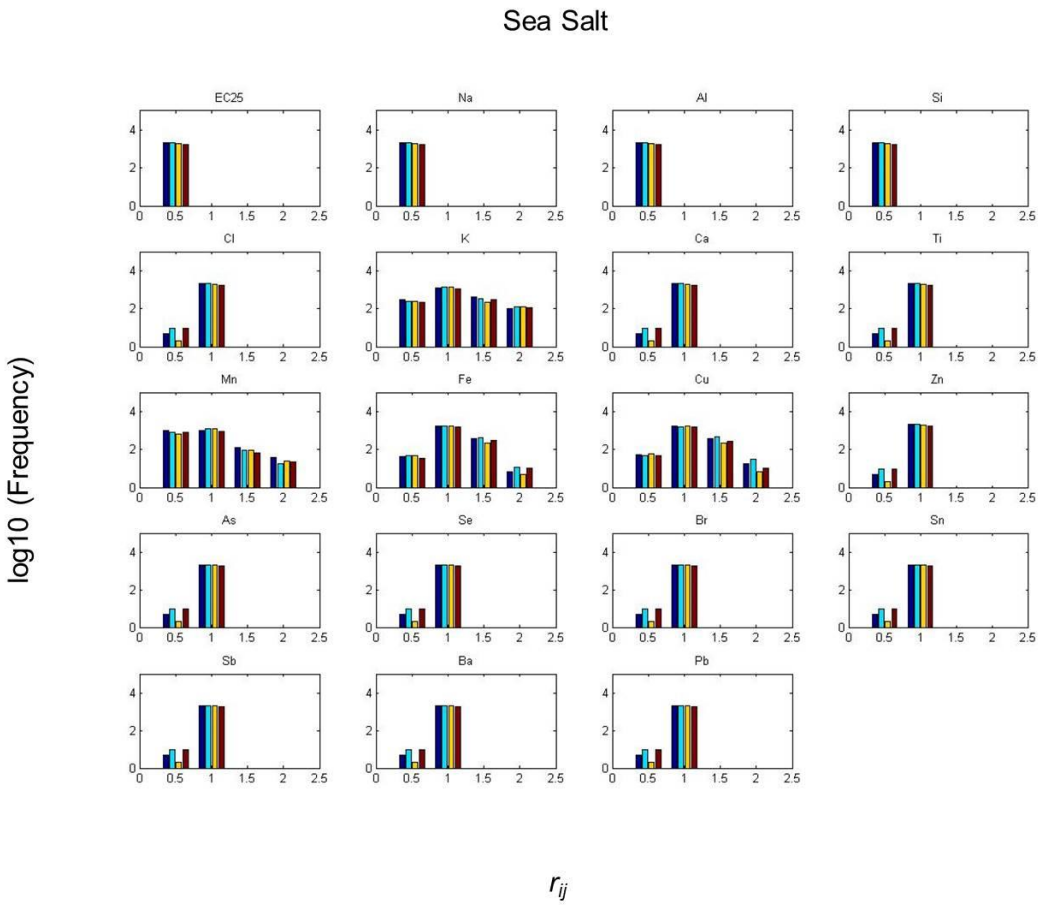

Figure C.19. Distribution of seasonally-stratified ratios rij for optimized sea salt source profiles for all available CSN monitors in 2006.

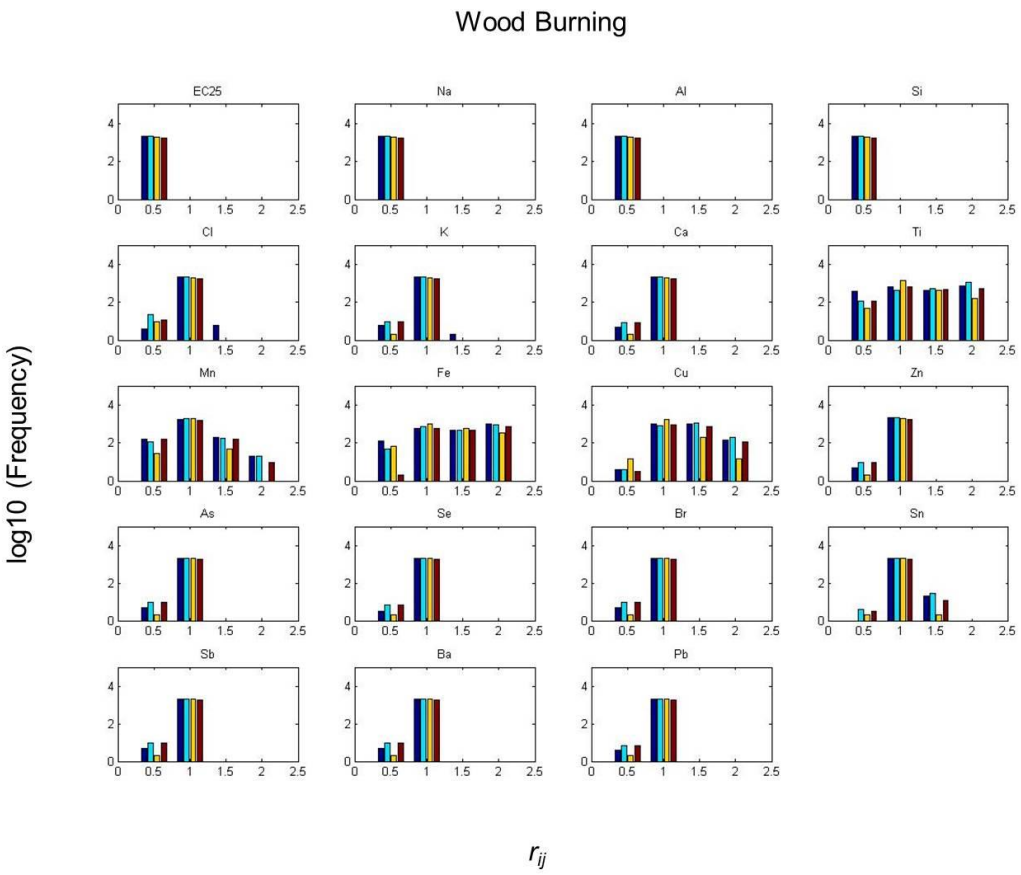

Figure C.20. Distribution of seasonally-stratified ratios rij for optimized wood burning source profiles for all available CSN monitors in 2006.

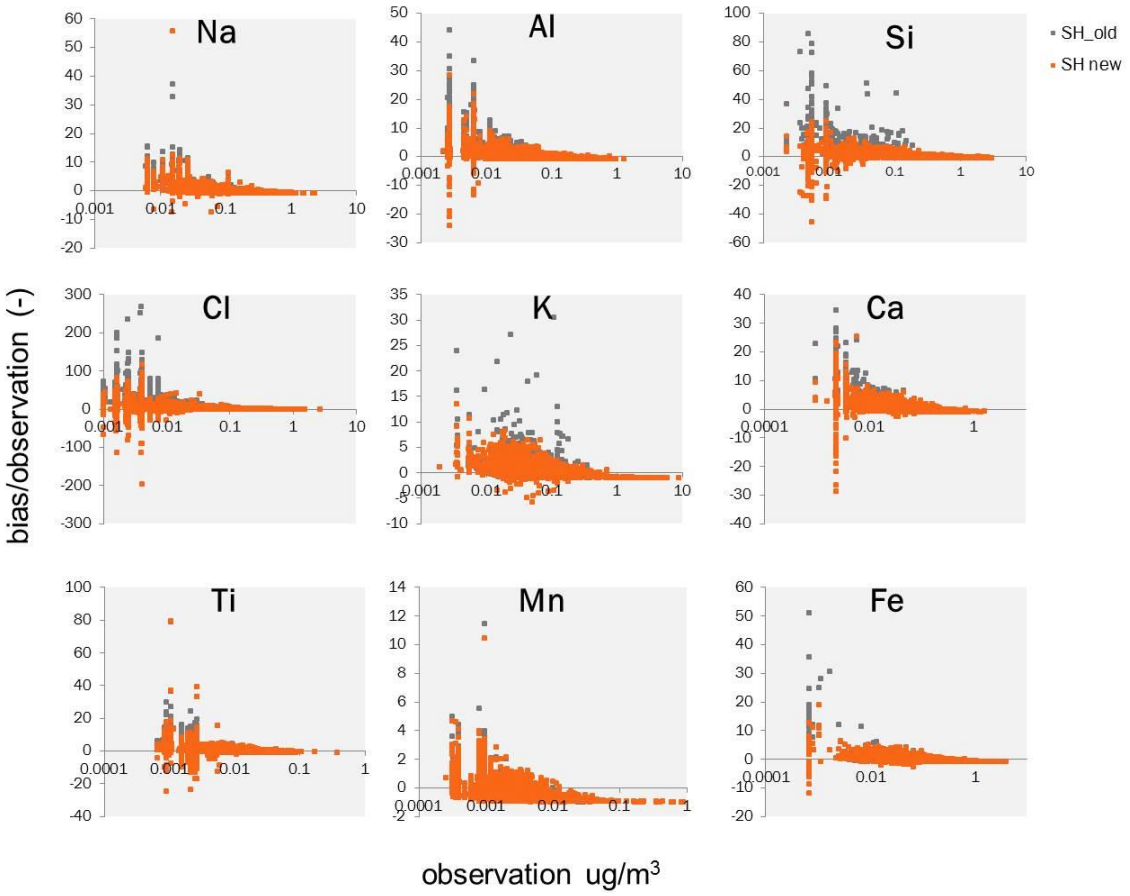

Figure C.21. Normalized bias vs. observations for metals concentrations that are reproduced using the optimized source profiles. Biases are normalized by observed concentration and are calculated for all available CSN site-days in 2006.

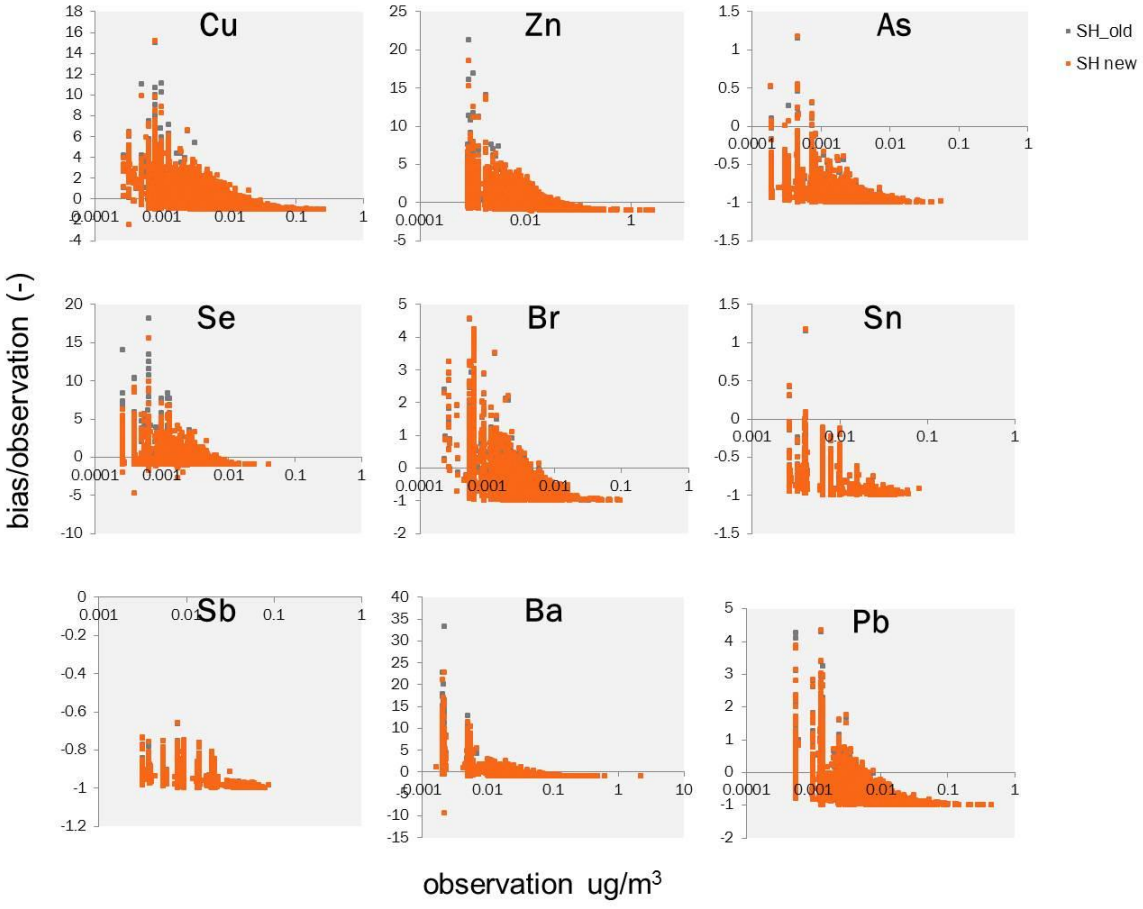

Figure C.21 Continued.

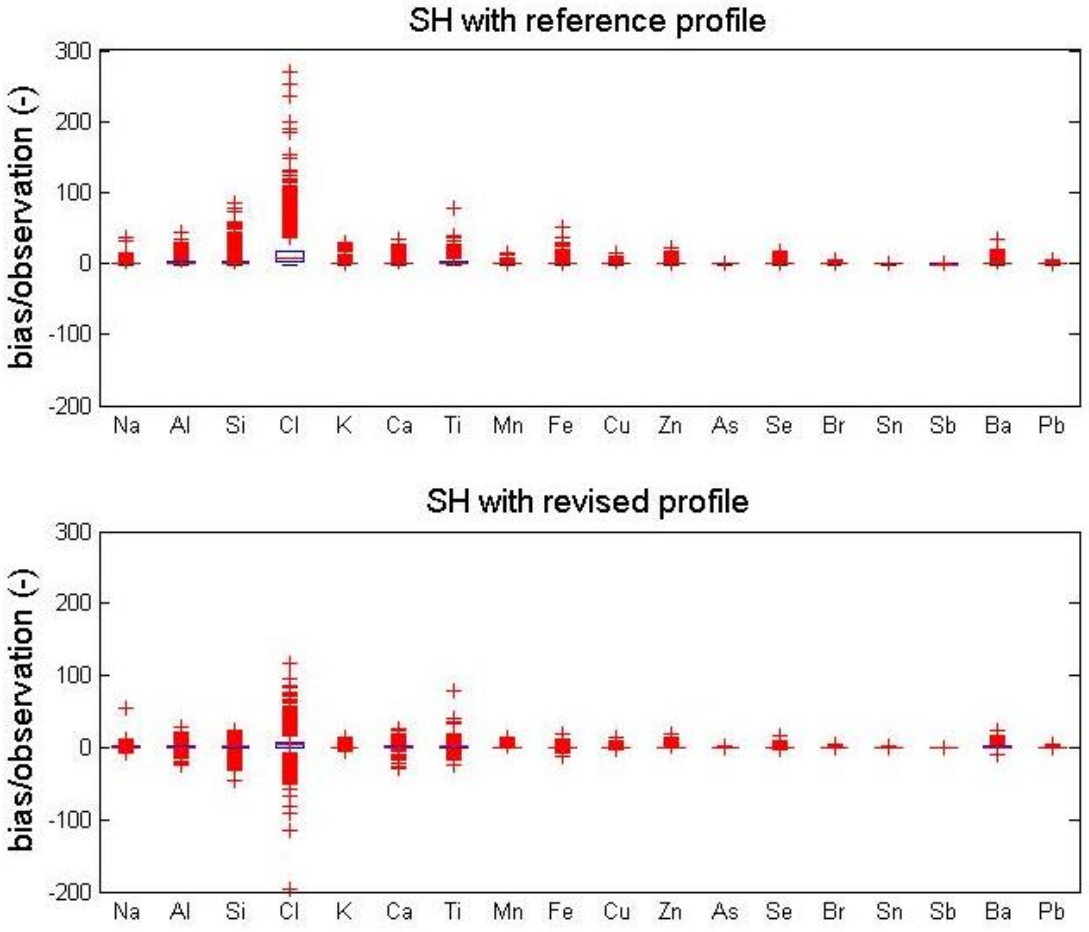

Figure C.22. Boxplots of normalized biases (unitless) for metals concentrations that are reproduced using the reference and optimized source profiles. Biases are normalized by observed concentration and are calculated for all available CSN site-days in 2006.

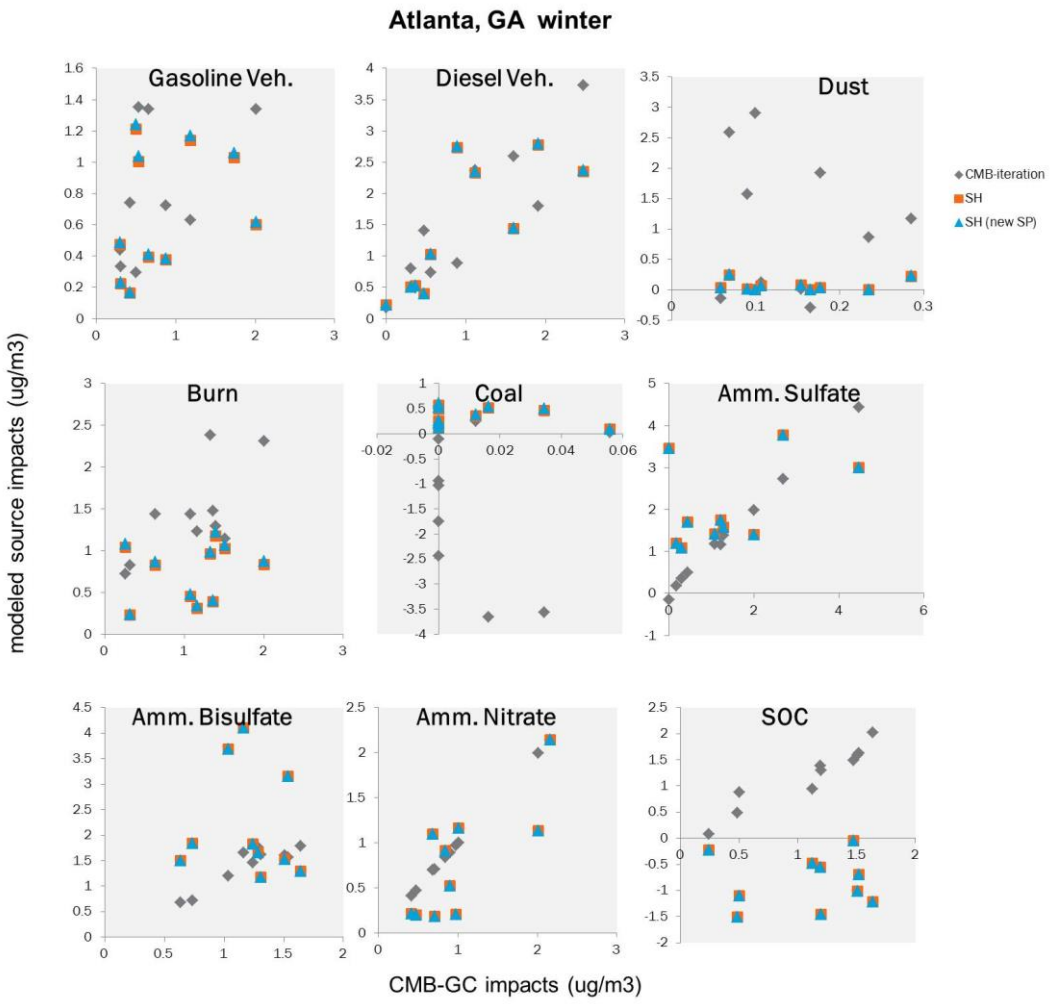

Figure C.23. Comparison of source impacts from CMB-GC with impacts from CMB-iteration with revised source profiles, SH impacts with reference profiles (SH), and SH impacts with revised profiles (SH new SP) for January 2006 in Atlanta, GA.

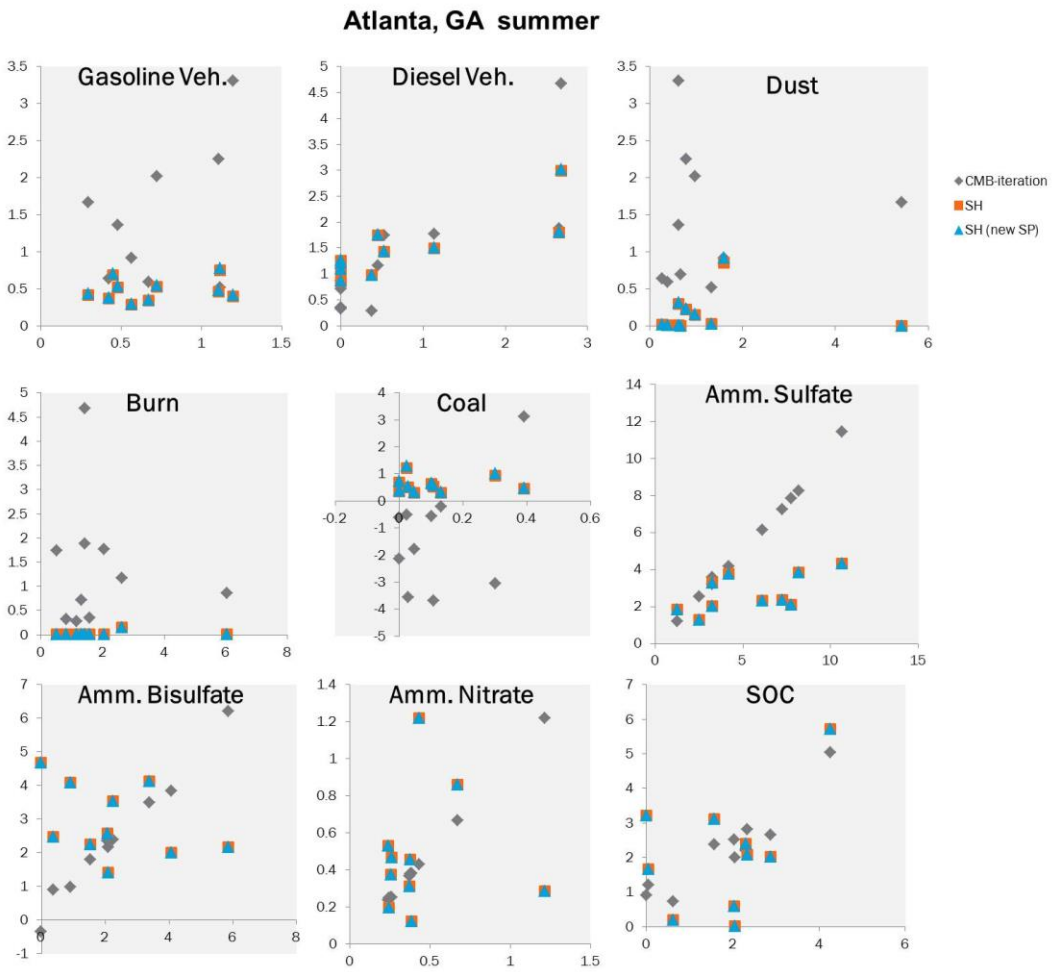

Figure C.24. Comparison of source impacts from CMB-GC with impacts from CMB-iteration with revised source profiles, SH impacts with reference profiles (SH), and SH impacts with revised profiles (SH new SP) for July 2006 in Atlanta, GA.

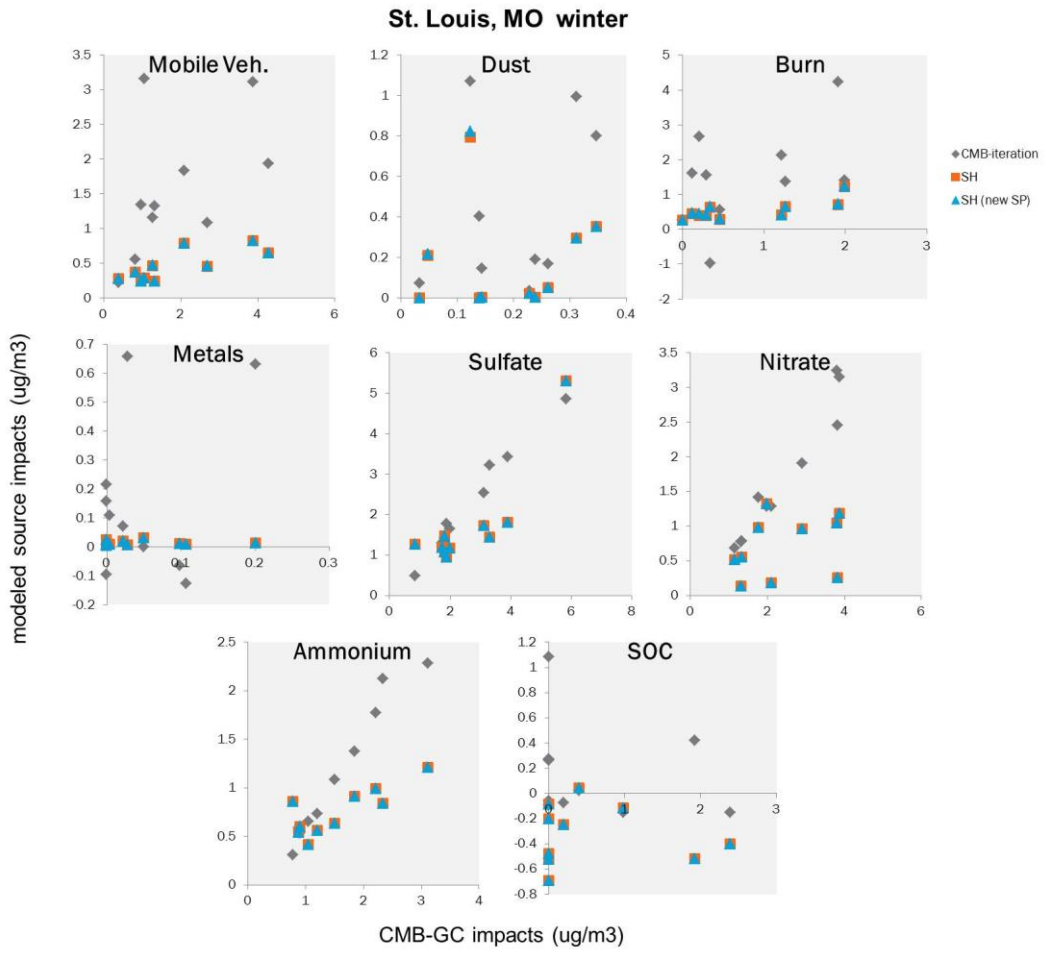

Figure C.25. Comparison of source impacts from CMB-GC with impacts from CMB-iteration with revised source profiles, SH impacts with reference profiles (SH), and SH impacts with revised profiles (SH new SP) for January 2006 in St. Louis, MO.

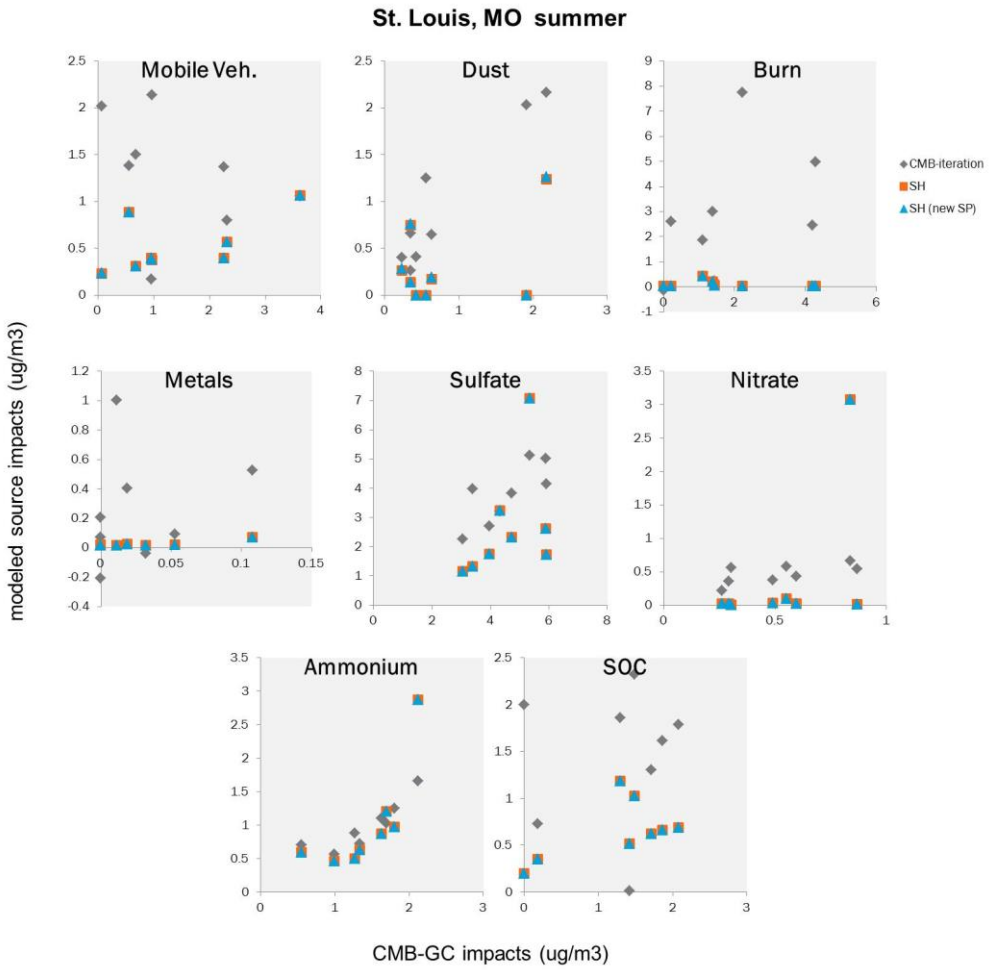

Figure C.26. Comparison of source impacts from CMB-GC with impacts from CMB-iteration with revised source profiles, SH impacts with reference profiles (SH), and SH impacts with revised profiles (SH new SP) for July 2006 in St. Louis, MO.

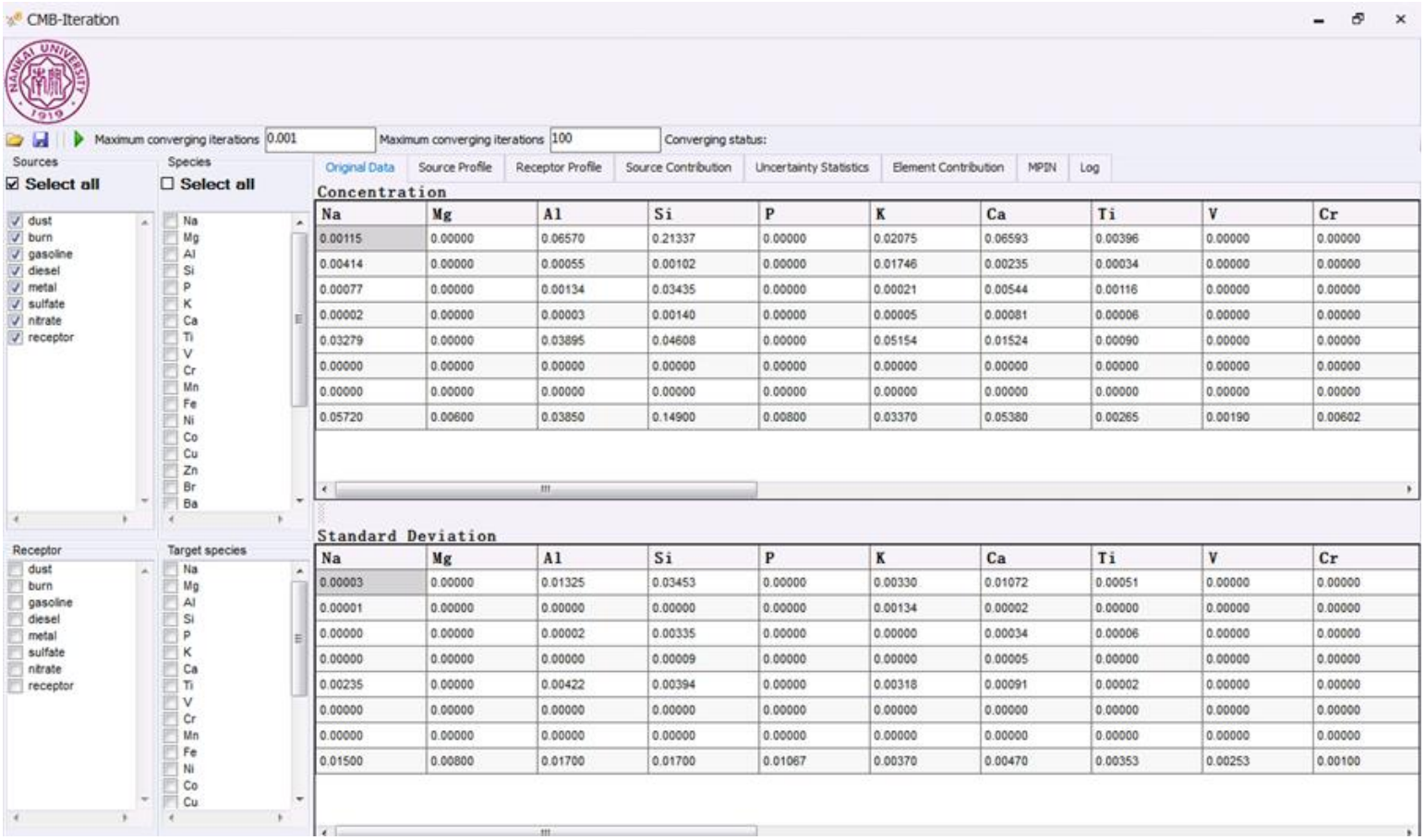

Figure C.27. Sample of the CMB-iteration (v3.0) user interface. The program is available through Nankai University (nksgl@nankai.edu.cn).

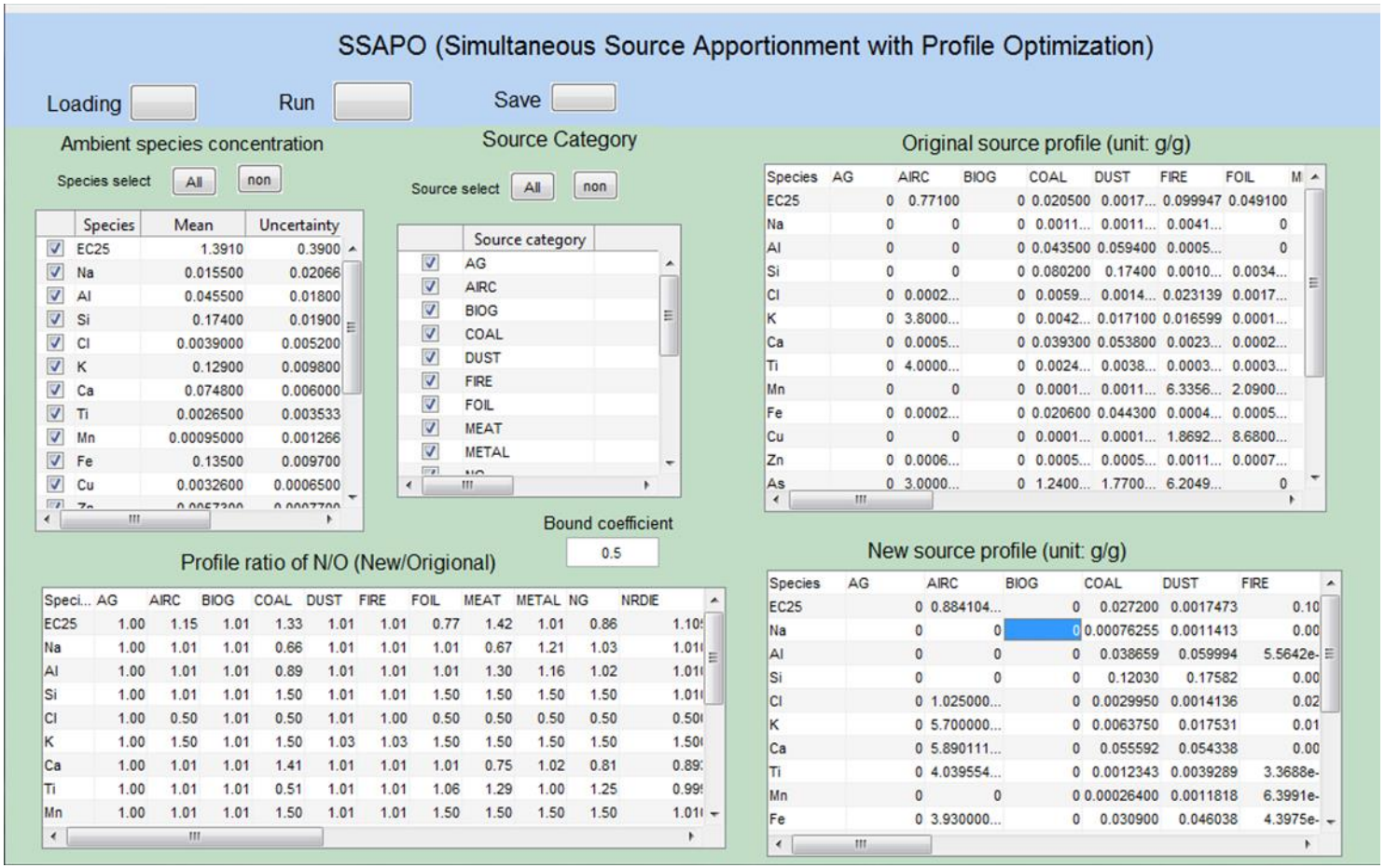

Figure C.28. Sample of the SSAPO user interface. The program is open-source and available through Georgia Institute of Technology (sunni.ivey@gmail.com).

APPENDIX D

CHAPTER 5 SUPPLEMENTAL INFORMATION

Table D.1: Source categories and abbreviations

source	abbrev.	source	abbrev.
agricultural activities and livestock operations	ag	non-road diesel	nrd
aircraft	air	non-road gasoline	nrg
biogenics	biog	non-road others	nro
coal combustion	coal	on-road diesel	ord
dust	dust	on-road gasoline	org
fires (wildfires, prescribed burns)	fire	others	ot
fuel oil combustion	foil	other combustion	otc
meat cooking	meat	solvents	slv
metals processing	metal	sea salt	ss
natural gas combustion	ng	wood burning	wood

Table D.2: Base model evaluation for CMAQ estimation of total $PM_{2.5}$, organic carbon, nitrate, ammonium, and sulfate concentrations (initial estimations prior to hybrid optimization). Comparison is conducted for all CSN sites-days for the year 2006. Concentrations are in units of $\mu g \cdot m^{-3}$.

metric	**$PM_{2.5}$**		**OC**		**NO3**		**NH4**		**SO4**	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
mean observation	12.6	7.63	2.22	1.76	1.58	2.29	1.40	1.18	3.13	2.61
mean CMAQ	12.6	8.58	1.60	1.27	1.76	2.55	1.47	1.22	2.89	2.30
mean bias	-0.001		-0.28		0.11		0.05		-0.08	
normalized mean error	0.46		0.53		0.79		0.49		0.38	

Table D.3: Mean (μ) and standard deviation (σ) of SC_{ij} values (μg·m^{-3}) for all CSN sites-days for the year 2006.

source	OC		NO3		NH4		SO4	
	μ	σ	μ	σ	μ	σ	μ	σ
ag	0.0031	0.020	0.084	0.81	0.038	0.31	0.00004	0.00075
air	0.00006	0.00069	0.00017	0.0057	0.00014	0.0038	0.00041	0.0019
biog	0.87	1.30	0.0012	0.067	-0.00017	0.014	-0.0001	0.0069
coal	0.0059	0.029	-0.0044	0.13	0.025	0.26	0.27	1.28
dust	0.00011	0.0017	0.0016	0.023	0.00079	0.0096	0.0010	0.0065
fire	0.00032	0.0081	0.00034	0.0056	0.00041	0.0091	0.00021	0.0034
foil	0.0017	0.0058	0.022	0.22	0.020	0.16	0.045	0.23
meat	0.00077	0.0052	0.0011	0.012	0.00050	0.0062	0.00024	0.0022
metal	0.00089	0.0028	0.00009	0.0069	0.0011	0.012	0.0085	0.053
ng	0.0015	0.0076	0.031	0.25	0.0092	0.078	0.0091	0.030
nrd	0.0016	0.011	0.012	0.17	0.0054	0.056	0.022	0.076
nrg	0.036	0.077	0.00037	0.049	-0.0002	0.012	0.00046	0.0067
nro	0.0015	0.0065	0.0075	0.062	0.0067	0.049	0.021	0.12
ord	0.0056	0.021	0.012	0.24	0.0054	0.067	0.023	0.076
org	0.026	0.063	0.079	0.52	0.022	0.16	0.016	0.053
ot	0.0011	0.013	0.016	0.13	0.012	0.071	0.041	0.14
otc	0.00067	0.0036	0.0031	0.025	0.0021	0.013	0.0073	0.031
slv	0.19	0.32	0.0028	0.039	0.00055	0.014	0.00068	0.0078
ss	0.00	0.00	0.00	0.00	0.00	0.00	0.0015	0.013
wood	0.019	0.16	0.0032	0.055	0.0011	0.015	0.0028	0.012

Table D.4: Mean (μ) and standard deviation (σ) of observed and modeled concentrations (μg·m^{-3}) for all CSN sites-days for the year 2006.

species	**$PM_{2.5}$**		**OC**		**NO3**		**NH4**		**SO4**	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
observation	12.6	7.63	2.22	1.76	1.58	2.29	1.40	1.18	3.13	2.61
CMAQ-DDM	12.6	8.57	1.61	1.27	1.76	2.55	1.47	1.22	2.89	2.30
hybrid	8.52	5.53	1.08	0.80	1.37	2.12	1.27	1.04	2.65	2.12
secondary adjusted	10.5	6.26	2.22	1.76	1.58	2.29	1.40	1.18	3.13	2.61

Table D.5: Statistics for 10-fold cross-validation at CSN sites. Comparison of original and withheld SC_{ij} values incudes the correlation coefficient (r), linear regression intercept (α), linear regression slope (β), and normalized mean bias (NMB). Regression is based on 7955 data pairs.

	OC				**NO3**				**NH4**				**SO4**			
	r	α	β	NMB	r	α	β	NMB	r	α	β	NMB	r	α	β	NMB
ag	0.76	0.00	0.86	-0.04	0.64	-0.02	0.82	-0.32	0.65	-0.01	0.77	-0.40	0.52	0.00	0.57	-0.45
air	0.06	0.00	0.09	0.00	0.32	0.00	0.54	-0.54	0.29	0.00	0.45	-0.47	0.41	0.00	0.53	-0.21
biog	0.62	0.17	0.81	0.00	0.30	0.00	0.41	-0.65	0.25	0.00	0.29	-1.55	0.04	0.00	0.06	2.32
coal	0.63	0.00	0.74	-0.05	0.33	0.00	0.57	-10.9	0.69	-0.01	0.88	-0.40	0.78	-0.06	0.91	-0.27
dust	0.23	0.00	0.34	0.01	0.49	0.00	1.10	0.31	0.31	0.00	0.58	0.14	0.47	0.00	1.01	0.43
fire	0.02	0.00	0.04	0.15	0.20	0.00	0.44	-0.08	0.07	0.00	0.16	-0.09	0.23	0.00	0.37	0.00
foil	0.66	0.00	0.84	-0.01	0.57	0.00	1.20	0.16	0.49	0.00	0.82	-0.06	0.52	0.01	0.68	-0.21
meat	0.52	0.00	0.71	-0.02	0.23	0.00	0.44	0.06	0.23	0.00	0.47	-0.04	0.39	0.00	0.60	-0.20
metal	0.53	0.00	0.74	0.03	0.36	0.00	0.58	-0.61	0.50	0.00	0.68	-0.34	0.54	0.00	0.61	-0.28
ng	0.43	0.00	0.59	-0.02	0.30	0.02	0.39	-0.18	0.28	0.00	0.36	-0.25	0.67	0.00	0.79	-0.19
nrd	0.37	0.00	0.50	-0.03	0.55	0.00	0.84	-0.27	0.52	0.00	0.77	-0.31	0.67	0.00	0.82	-0.08
nrg	0.50	0.01	0.69	0.03	0.30	0.00	0.61	-0.78	0.27	0.00	0.50	-1.90	0.65	0.00	0.89	-0.31
nro	0.44	0.00	0.68	0.04	0.43	0.00	0.63	-0.22	0.35	0.00	0.47	-0.23	0.53	0.00	0.66	-0.18
ord	0.46	0.00	0.63	0.01	0.54	0.00	0.73	-0.46	0.54	0.00	0.69	-0.45	0.76	0.00	0.83	-0.19
org	0.46	0.01	0.67	0.04	0.52	0.01	0.67	-0.21	0.52	0.00	0.63	-0.32	0.78	0.00	0.86	-0.14
ot	0.55	0.00	0.95	0.08	0.50	0.00	0.80	0.02	0.49	0.00	0.77	0.00	0.57	0.01	0.94	0.14
otc	0.39	0.00	0.49	-0.03	0.39	0.00	0.52	-0.25	0.40	0.00	0.49	-0.30	0.47	0.00	0.57	-0.25
slv	0.54	0.06	0.70	0.00	0.37	0.00	0.53	-0.21	0.31	0.00	0.47	-0.30	0.58	0.00	0.70	-0.28
ss	-	-	-	-	-	-	-	-	-	-	-	-	0.51	0.00	0.68	-0.15
wood	0.11	0.02	0.19	0.04	0.22	0.00	0.67	0.04	0.22	0.00	0.59	-0.08	0.49	0.00	0.74	-0.05

Table D.6: Evaluation metrics for withheld site-days in the 10-fold cross-validation at CSN sites. Presented is a comparison of observations and concentrations that were calculated using withheld (wh) and kriged (krig) SCij values. Evaluation metrics include the correlation coefficient (r), linear regression intercept (α), linear regression slope (β), and normalized mean bias (NMB). Regression is based on 7955 data pairs.

	OC		**NO3**		**NH4**		**SO4**	
	obs vs. wh	obs vs. krig	obs vs. wh	obs vs. krig	obs vs. wh	obs vs. krig	obs vs. wh	obs vs. krig
r	1.00	0.69	1.00	0.70	1.00	0.79	1.00	0.91
α	0.00	0.43	0.00	0.39	0.00	0.22	0.00	0.22
β	1.00	0.81	1.00	0.72	1.00	0.81	1.00	0.89
NMB	0.00	0.00	0.00	0.06	0.00	0.05	0.00	0.04

Table D.7: Seasonal observations, secondary-adjusted concentrations, and source impacts for total $PM_{2.5}$ for 2006 for Atlanta, GA. Concentrations and source impacts are in units of $\mu g \cdot m^{-3}$.

species	**winter ($\mu g \cdot m^{-3}$)**	**spring ($\mu g \cdot m^{-3}$)**	**summer ($\mu g \cdot m^{-3}$)**	**fall ($\mu g \cdot m^{-3}$)**
observation	11.8	15.0	22.3	15.3
secondary-adjusted conc.	14.1	15.7	20.1	15.6
ag	1.47	1.26	0.78	1.39
air	0.23	0.20	0.36	0.26
biog	1.27	1.35	2.78	1.67
coal	1.12	3.36	6.32	3.22
dust	0.058	0.11	0.25	0.084
fire	0.019	0.033	0.010	0.0066
foil	0.17	0.40	0.63	0.35
meat	0.26	0.23	0.30	0.27
metal	0.024	0.048	0.072	0.040
ng	0.07	0.14	0.18	0.12
nrd	0.61	0.70	1.04	0.79
nrg	0.19	0.20	0.31	0.26
nro	0.039	0.11	0.10	0.066
ord	1.65	1.54	2.35	1.76
org	1.85	1.11	1.51	1.45
ot	0.81	0.99	1.09	0.75
otc	0.023	0.047	0.068	0.041
slv	0.35	0.13	0.12	0.24
ss	0.029	0.022	0.0034	0.014
wood	0.25	0.30	0.041	0.18
sum of impacts	10.5	12.3	18.3	13.0

Table D.8: Seasonal observations, secondary-adjusted concentrations, and source impacts for total $PM_{2.5}$ for 2006 for Los Angeles, CA. Concentrations and source impacts are in units of $\mu g \cdot m^{-3}$.

species	**winter ($\mu g \cdot m^{-3}$)**	**spring ($\mu g \cdot m^{-3}$)**	**summer ($\mu g \cdot m^{-3}$)**	**fall ($\mu g \cdot m^{-3}$)**
observation	17.2	12.8	16.8	15.2
secondary-adjusted conc.	18.4	17.5	20.5	18.5
ag	1.33	1.80	0.84	1.16
air	0.014	0.026	0.009	0.008
biog	2.64	0.85	1.05	1.84
coal	0.41	0.21	0.31	0.23
dust	0.17	0.19	0.18	0.13
fire	0.029	0.042	0.052	0.032
foil	2.07	1.21	1.09	1.36
meat	0.57	0.68	0.60	0.61
metal	0.11	0.10	0.14	0.11
ng	1.35	1.84	1.33	1.27
nrd	1.45	2.10	2.40	2.08
nrg	0.20	0.24	0.35	0.26
nro	0.16	0.79	1.40	0.50
ord	1.78	1.32	1.15	1.46
org	3.62	3.57	3.08	3.05
ot	0.73	0.88	1.00	0.72
otc	0.23	0.31	0.32	0.25
slv	0.83	0.41	0.23	0.36
ss	0.033	0.13	0.16	0.077
wood	0.43	0.24	0.010	0.27
sum of impacts	18.2	16.9	15.7	15.8

Table D.9: Seasonal observations, secondary-adjusted concentrations, and source impacts for total $PM_{2.5}$ for 2006 for Pittsburgh, PA. Concentrations and source impacts are in units of $\mu g \cdot m^{-3}$.

species	**winter ($\mu g \cdot m^{-3}$)**	**spring ($\mu g \cdot m^{-3}$)**	**summer ($\mu g \cdot m^{-3}$)**	**fall ($\mu g \cdot m^{-3}$)**
observation	12.5	13.3	20.6	12.9
secondary-adjusted conc.	11.8	13.4	18.9	13.7
ag	1.95	1.67	1.38	1.64
air	0.021	0.023	0.034	0.022
biog	0.40	0.36	0.13	0.62
coal	1.25	2.53	6.94	2.16
dust	0.042	0.10	0.083	0.052
fire	0.0069	0.0042	0.0023	0.0067
foil	0.20	0.21	0.31	0.19
meat	0.25	0.34	0.37	0.33
metal	0.27	0.28	0.43	0.28
ng	0.13	0.12	0.24	0.10
nrd	0.66	0.96	1.66	0.99
nrg	0.16	0.16	0.21	0.21
nro	0.015	0.042	0.083	0.029
ord	0.52	0.62	1.07	0.62
org	1.36	1.04	1.19	1.17
ot	1.06	1.17	1.33	1.01
otc	0.032	0.053	0.12	0.038
slv	0.30	0.18	0.087	0.31
ss	0.0043	0.0048	0.0036	0.0058
wood	0.26	0.36	0.039	0.24
sum of impacts	8.89	10.2	15.7	10.0

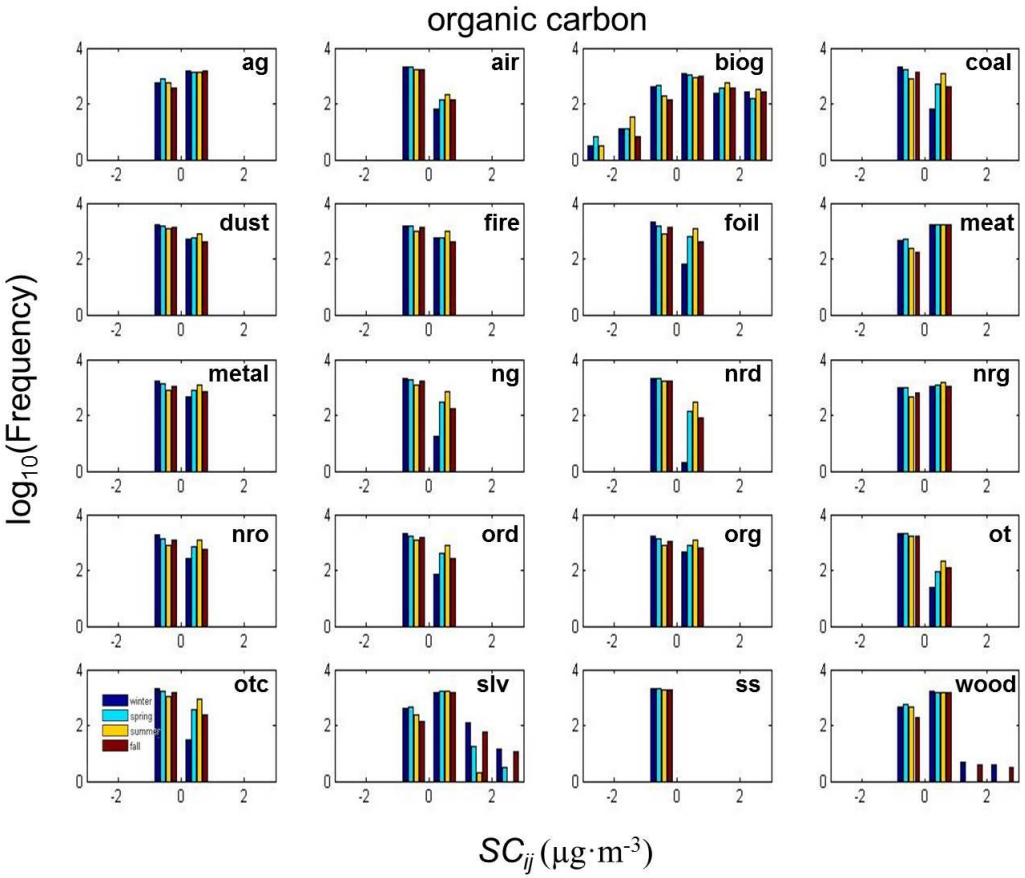

Figure D.1: Distributions of SC_{ij} ($\mu g \cdot m^{-3}$) at CSN sites for organic carbon in 2006, stratified by seasons: winter (dark blue), spring (cyan), summer (yellow), fall (dark red).

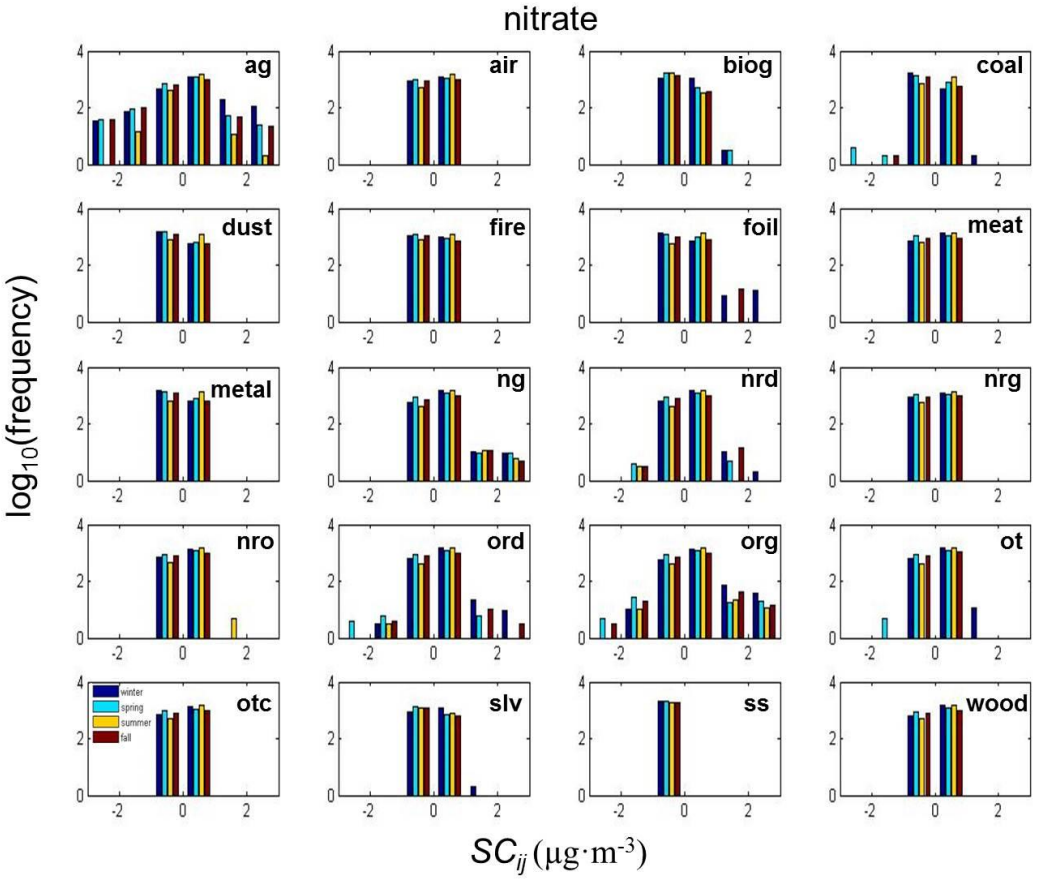

Figure D.2: Distributions of SC_{ij} ($\mu g \cdot m^{-3}$) at CSN sites for nitrate in 2006, stratified by seasons: winter (dark blue), spring (cyan), summer (yellow), fall (dark red).

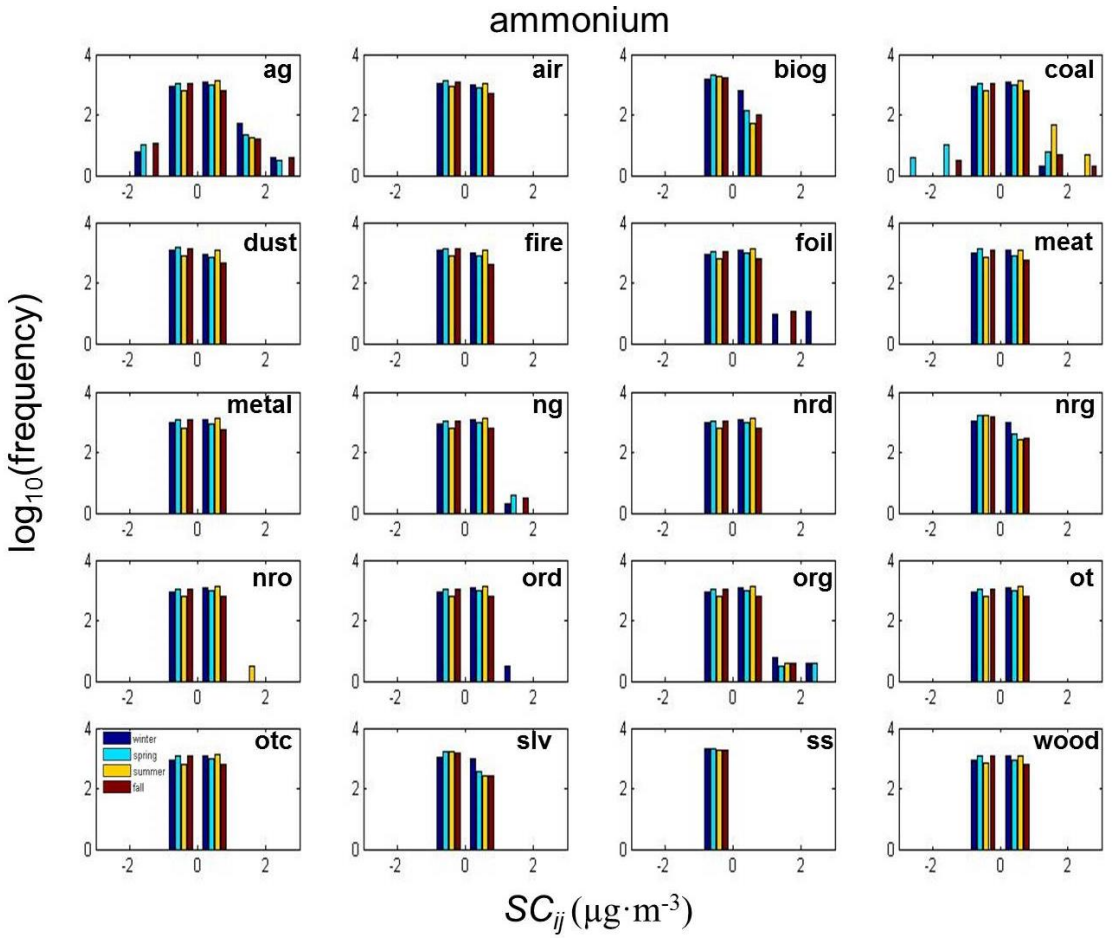

Figure D.3: Distributions of SC_{ij} ($\mu g \cdot m^{-3}$) at CSN sites for ammonium in 2006, stratified by seasons: winter (dark blue), spring (cyan), summer (yellow), fall (dark red).

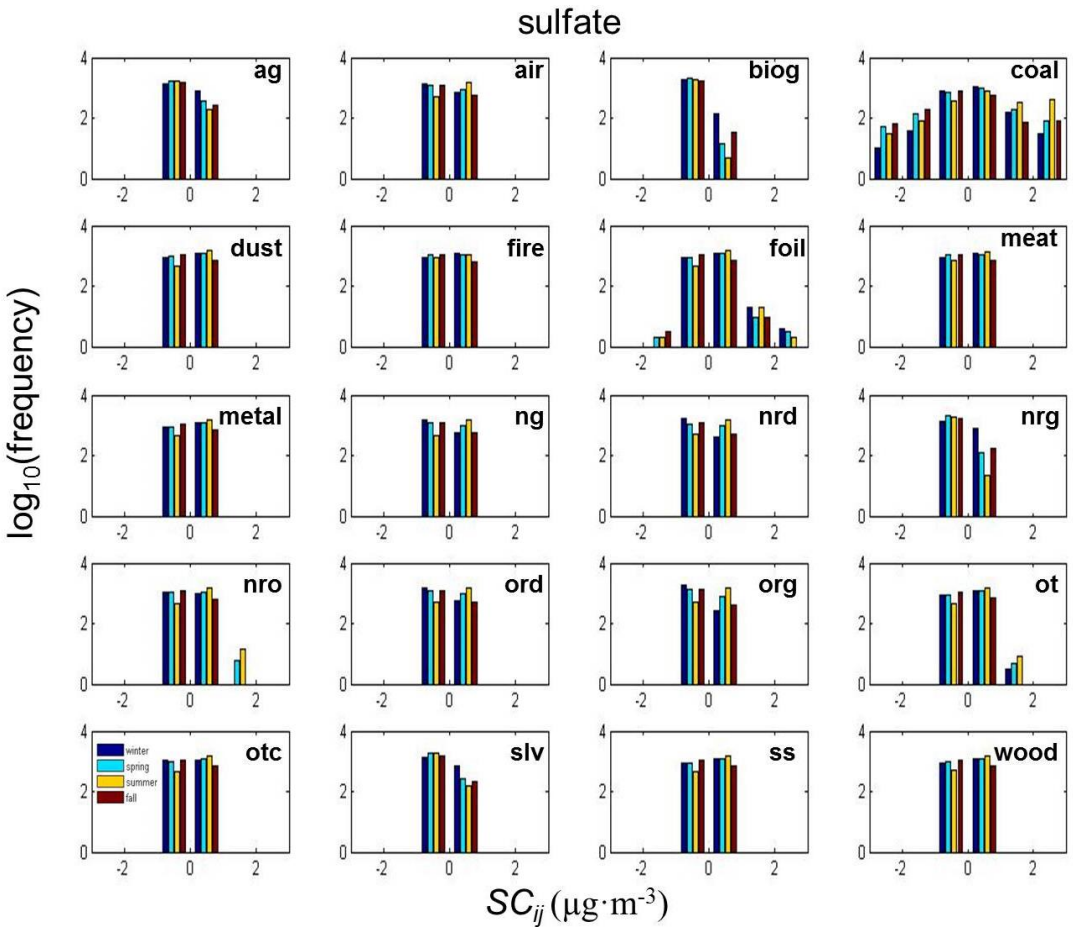

Figure D.4: Distributions of SC_{ij} ($\mu g \cdot m^{-3}$) at CSN sites for sulfate in 2006, stratified by seasons: winter (dark blue), spring (cyan), summer (yellow), fall (dark red).

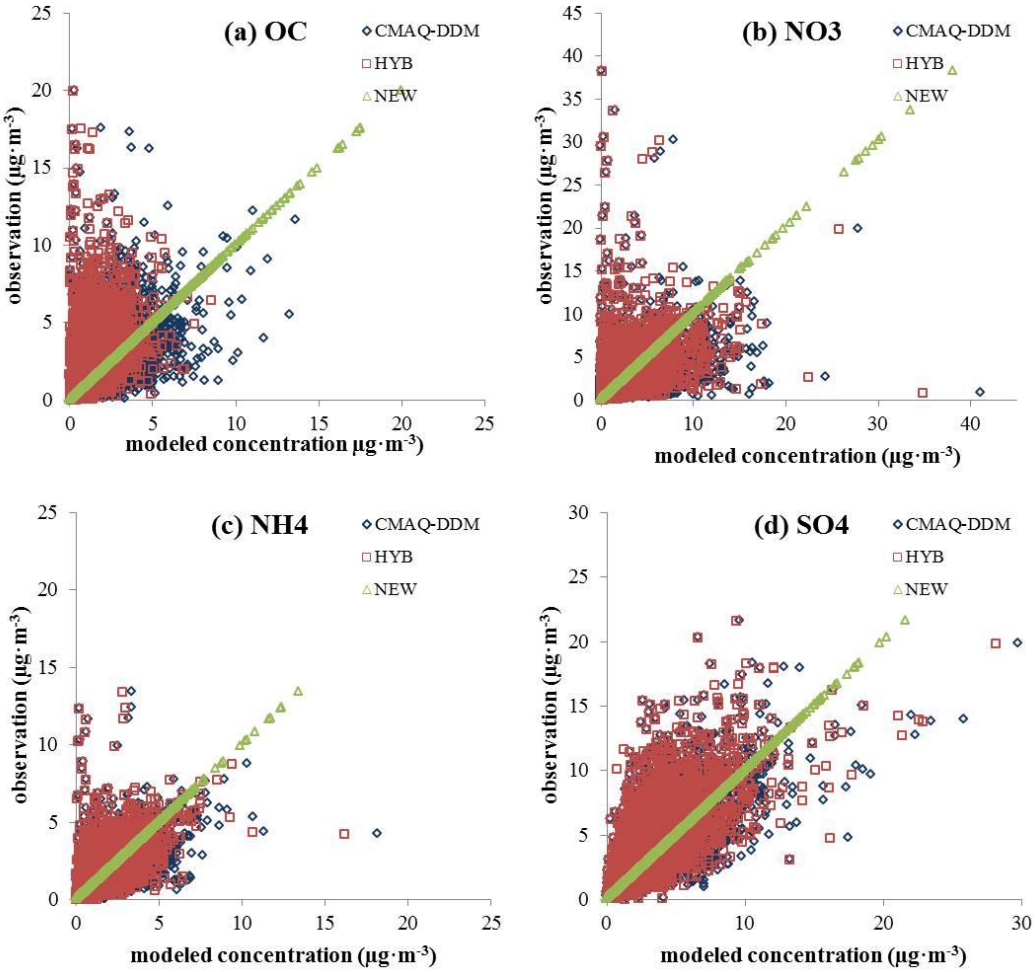

Figure D.5: Observations vs. simulated concentrations of (a) organic carbon, (b) nitrate, (c) ammonium, and (d) sulfate at CSN locations in 2006. Modeled concentrations include original CMAQ-DDM, hybrid (HYB), and the secondary-adjusted (NEW) concentrations

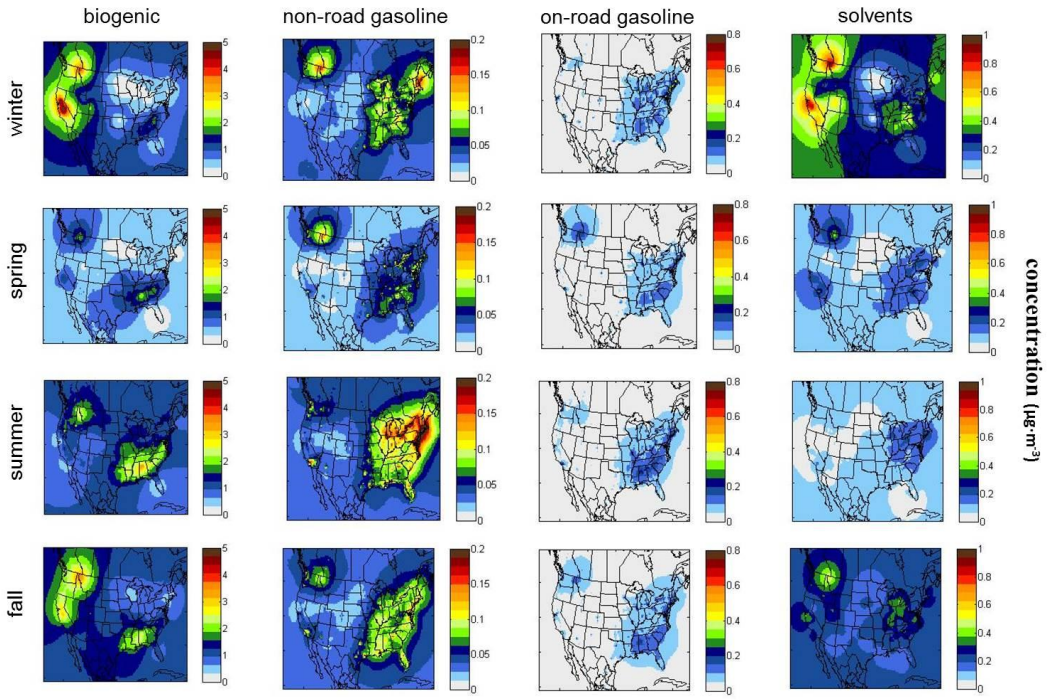

Figure D.6: Seasonally-averaged spatial fields of secondary-adjusted source impacts in $\mu g \cdot m^{-3}$ for organic carbon and biogenics (column 1), non-road gasoline (column 2), on-road gasoline (column 3), and solvents (column 4).

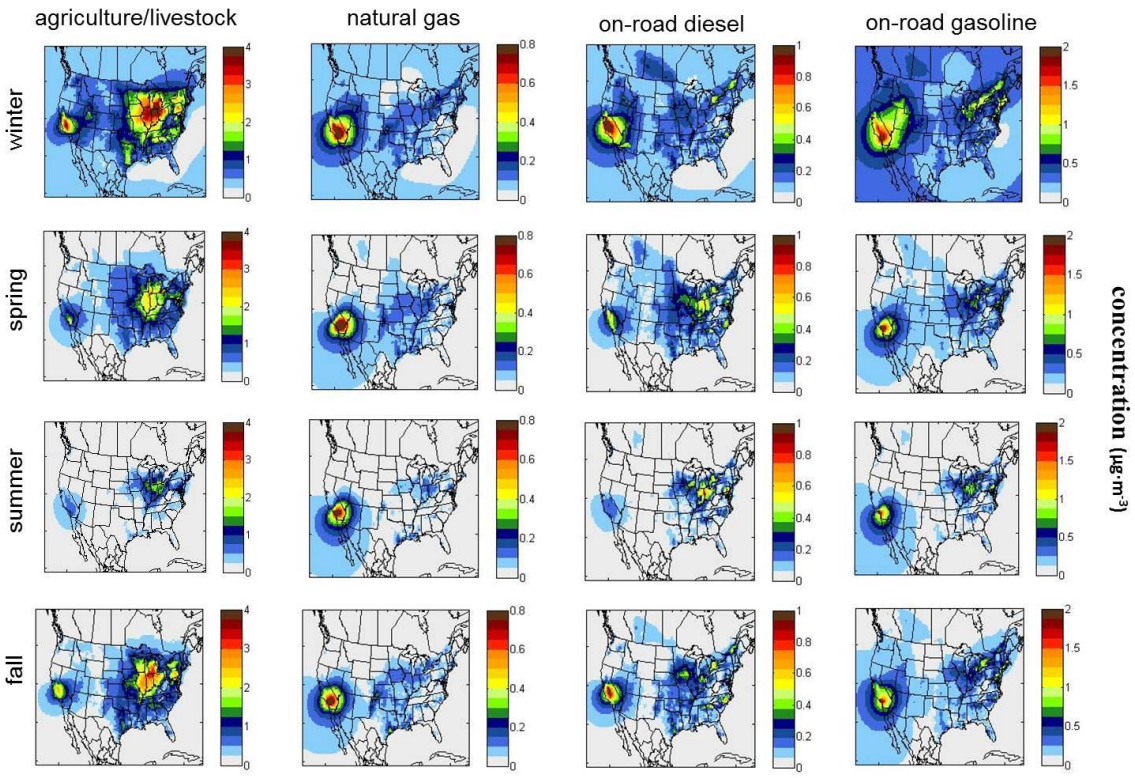

Figure D.7: Seasonally-averaged spatial fields of secondary-adjusted source impacts in $\mu g \cdot m^{-3}$ for nitrate and agriculture/livestock (column 1), natural gas combustion (column 2), on-road diesel (column 3), and on-road gasoline (column 4).

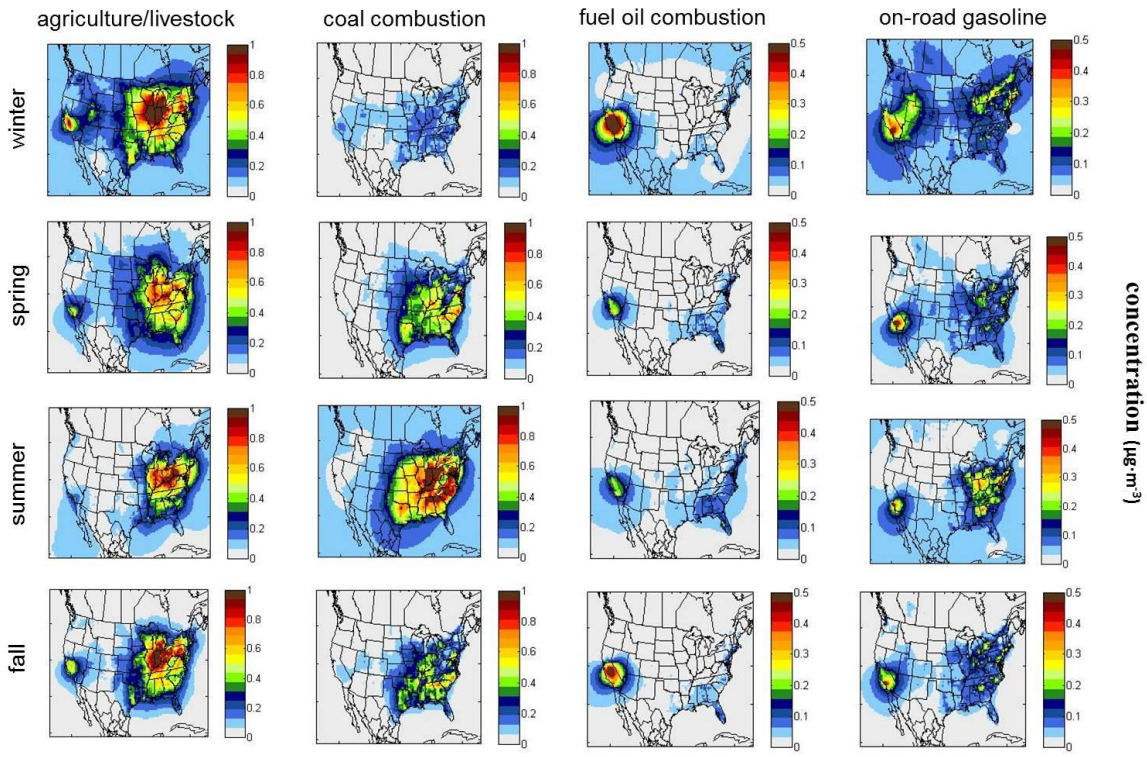

Figure D.8: Seasonally-averaged spatial fields of secondary-adjusted source impacts in $\mu g \cdot m^{-3}$ for ammonium and agricultural activities (column 1), coal combustion (column 2), fuel oil combustion (column 3), and on-road gasoline (column 4).

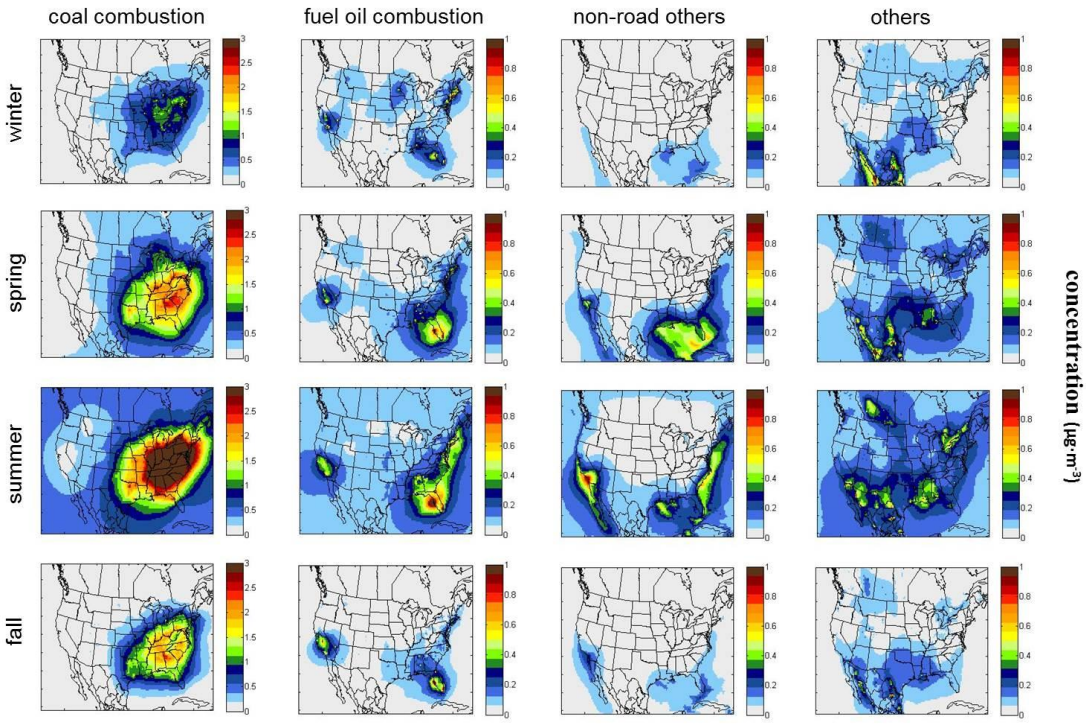

Figure D.9: Seasonally-averaged spatial fields of secondary-adjusted source impacts in $\mu g \cdot m^{-3}$ for sulfate and coal combustion (column 1), fuel oil combustion (column 2), non-road others (column 3), and other $PM_{2.5}$ sources (column 4).

VITA

CESUNICA E. IVEY

Cesunica E. Ivey was born in Atlanta, GA and graduated from Westlake High School (Atlanta, GA) as valedictorian in 2006. She received a B.S. in Mathematics from Fort Valley State University (Fort Valley, GA) in 2008, a B.S. in Civil Engineering from Georgia Institute of Technology (Atlanta, GA) in 2010, and an M.S. in Environmental Engineering from Georgia Institute of Technology (Atlanta, GA) in 2011. When she is not conducting research, Cesunica spends her time caring for her brother and nieces, babysitting for local Atlanta families, mentoring other doctoral students, and indulging in cinema and athletics.